PCCP

Cite this: *Phys. Chem. Chem. Phys.*, 2012, **14**, 7809–7820

www.rsc.org/pccp

PAPER

Extended implementation of canonical transformation theory: parallelization and a new level-shifted condition

Takeshi Yanai,*[a] Yuki Kurashige,[a] Eric Neuscamman†[b] and Garnet Kin-Lic Chan[b]

Received 28th November 2011, Accepted 16th April 2012
DOI: 10.1039/c2cp23767a

The canonical transformation (CT) theory has been developed as a multireference electronic structure method to compute high-level dynamic correlation on top of a large active space reference treated with the *ab initio* density matrix renormalization group method. This article describes a parallelized algorithm and implementation of the CT theory to handle large computational demands of the CT calculation, which has the same scaling as the coupled cluster singles and doubles theory. To stabilize the iterative solution of the CT method, a modification to the CT amplitude equation is introduced with the inclusion of a level shift parameter. The level-shifted condition has been found to effectively remove a type of intruder state that arises in the linear equations of CT and to address the discontinuity problems in the potential energy curves observed in the previous CT studies.

1 Introduction

Dynamic correlation is a key description in multireference electronic structure calculations to deliver quantitative accuracy to the active space description whose accuracy is at a qualitative level with static correlation alone.[1–71] The way of viewing electron correlation in terms of static and dynamic correlations is a well-established concept in the multireference theory, and the active space model that treats these correlations on the separated physical scales is one of the most successful approaches. The active space is selected semi-manually so as to provide a qualitatively good approximation to the exact solution modeled by full configuration interaction (FCI). This leads to the following expansion of the FCI wavefunction:

$$|\Psi_{\mathrm{FCI}}\rangle = |\Psi_{\mathrm{act}}\rangle + |\Psi_{\Delta}\rangle, \quad (1)$$

where $|\Psi_{\mathrm{act}}\rangle$ is the active space wavefunction that is generally multiconfigurational and describes static correlation. Eqn (1) is considered to be heavily weighted towards $|\Psi_{\mathrm{act}}\rangle$, and the rest of the expansion $|\Psi_{\Delta}\rangle$ is a perturbative or small residual, corresponding to dynamic correlation.

One of the well-established prescriptions to construct the multiconfigurational wavefunction for $|\Psi_{\mathrm{act}}\rangle$ is the complete active space (CAS) approach, developed by Roos *et al.*,[72,73] equivalently by Ruedenberg *et al.*[74] (with the different name *fully optimized reaction space*). The CAS model has several advantages in terms of clearness of specifying active space, size-consistency, *etc.* Since $|\Psi_{\mathrm{act}}\rangle$ for CAS is traditionally determined by CAS-CI calculation, namely FCI diagonalization of the active space Hamiltonian, the obvious problem arises from the exponential dependence of the CAS-CI algorithm on the size of active space.

Recent works of our group or others have approached the complexity of active space correlation by using the density matrix renormalization group (DMRG) method[75,76] as a direct substitute for CASCI (or FCI).[77–106] The static correlation is thought of as a manifestation of the inter-atomic overlap of near-degenerate valence atomic states, and thus can be efficiently, adaptively described by the local multireference structure of the DMRG wavefunction. Orbital optimization coupled with the active space DMRG calculations in a self-consistent field (SCF) manner was introduced in ref. 97–100 and leads to the DMRG-SCF or DMRG-CASSCF methods. These models have been successfully applied to the covalent excited states of β-carotene[99] and the spin states of *m*-phenylenecarbene,[105] where unprecedent large-size active spaces, such as CAS(50e,50o), correlating 50 electrons within 50 active orbitals, were accurately handled with optimized orbitals. This extensibility is associated with the DMRG ansatz that is built upon an entanglement of the local interacting objects, as related to the theme of the present Special Issue.

The development of efficient multireference methods to calculate the dynamic correlation in conjunction with the active space description remains a challenging topic. This type of correlation, referred to as "multireference dynamic correlation," should be treated with the low-order many-body theories of weak correlation, such as perturbation theory (PT),[1–31] configuration interaction (CI),[32–49] or coupled cluster (CC) theory.[50–71]

[a] *Department of Theoretical and Computational Molecular Science, Institute for Molecular Science, Okazaki, Aichi, 444-8585, Japan. E-mail: yanait@ims.ac.jp*
[b] *Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York, 14853-1301, USA*
† Present address: Department of Chemistry, University of California, Berkeley, Berkeley California 94720.

In previous papers,[107,108] we showed the combination of a large-active-space DMRG treatment of static correlation and an exponential based treatment of dynamical correlation through the canonical transformation (CT) theory. The developments of CT theory were reported in ref. 107–114. It uses a canonical (*i.e.* unitary) exponential ansatz, is size-consistent, and thus may be considered a kind of multi-reference coupled-cluster (MRCC) theory. In the CT model, an emphasis is placed on an effective Hamiltonian picture of the dynamic correlation. The complexity of the exponential operator of the dynamic correlation is transferred from the wavefunction to the Hamiltonian in a way to avoid a direct manipulation of the complex reference wavefunction (*e.g.* DMRG wavefunction). The effective CT Hamiltonian is constructed approximately as a two-body description using an operator decomposition based on Mukherjee–Kutzelnigg normal ordering and density matrix cumulant decomposition[115–117] to achieve a higher-order, size-consistent treatment of dynamic correlation in a computationally efficient way. For the construction, the static correlation in the reference is taken into account using only the one- and two-body reduced density matrices (RDMs). This reduced reference treatment is alternatively referred to as the internally-contracted (IC) multi-reference algorithm, which was first introduced by Meyer,[36] was practically used in IC-MRCI by Werner *et al.*[45,46] (with partial uncontraction) as well as CASPT2 by Roos *et al.*,[1–3] and was recently investigated for the developments of IC-MRCC.[68–71] The CT theory exhibits accuracy on a par with the best MRCI approaches, but shares the same favorable sixth-power computational scaling as the single-reference coupled cluster theory. The quantum chemical applications of the joint CT and DMRG theory were shown in the copper-oxo dimer isomerization problem,[107] as well as in the study of excited states in porphin.[108] With the DMRG, we handled large active spaces (*e.g.* CAS(28e, 32o) and CAS(24e,24o) for Cu_2O_2 and porphin, respectively), and the remaining orbital correlation was incorporated through CT theory. It should be mentioned that two of the authors recently reported a combination of the DMRG-CASSCF and CASPT2 methods with the use of the three-body and contracted four-body RDMs of active space.[31]

In this paper, we present a detailed description of the parallelized implementation of the CT theory. The accurate evaluation of dynamic correlation requires a large basis representation, including polarization functions and in some cases diffuse Rydberg-like functions. The CT algorithm has a steep dependence of the computational demanding on the size of basis sets: $O(n_a^2\ N^4)$ and $O(N^4)$ for operation counts and memory storage, respectively, where n_a and N refer to the number of active orbitals and all the correlated orbitals, respectively. We attempt to overcome this computational difficulty by the parallelization method that distributes computational efforts and data across network-connected computers.

In addition, we will show an extension of CT theory with introduction of a level shift to the stationary condition for CT solutions. The central underlying numerical difficulties in the CT calculations are intruder states, which arise from the cumulant and operator decomposition approximations (see ref. 109–113), both making the CT equations too poorly conditioned. Previously, we developed two approaches to circumvent the intruder states: (1) the overlap truncation method,[110] aggressively eliminating linear dependence from the first order interacting basis through orthogonalization that involves $O(n_a^9)$ cost diagonalization, and (2) the use of strongly contracted excitation operators,[108] which was first introduced by Malrieu *et al.* in *n*-electron valence perturbation (NEVPT2) theory,[27–30] and intelligently restricts the first order basis to a single linear combination of active states for a given set of external orbitals. In this study, we propose an alternative approach which makes a radical change to the CT stationary equation so as to regularize its singularity. When this approach is used together with the orthogonalized or strongly contracted basis operators, the aggressive truncation in the construction of this basis can be avoided. This leads to a remediation in the undesirable feature of the CT method that yields nonsmooth (stepwise) potential curves.[109,110,112,113]

This paper proceeds as follows. In Section 2, a brief review of the CT method is followed by the introduction of the level-shifted stationary condition. Section 3 shows the details of the parallel algorithm for the CT theory along with its tensor contraction expressions. In Section 4, illustrative calculations are shown. We finish then with our summary.

2 Algorithm

2.1 Canonical transformation

Our canonical transformation (CT) theory[107–114] claims that dynamic correlation is described by a similarity transformation of bare Hamiltonian $\hat{H}$, leading to effective Hamiltonian $\hat{\bar{H}}$, as given by

$$\hat{\bar{H}} = e^{\hat{A}^\dagger} \hat{H} e^{\hat{A}}, \tag{2}$$

where the many-body operator $e^{\hat{A}}$ is set to be unitary with the excitation amplitude $\hat{A} = -\hat{A}^\dagger$. In second quantization with given orbital basis $\{\phi_p(r)\}$, $\hat{H}$ is expressed as,

$$\hat{H} = h_0 + \hat{h}_1 + \hat{h}_2, \tag{3a}$$

$$\hat{h}_1 = t_{q_1}^{p_1} \hat{E}_{q_1}^{p_1}, \tag{3b}$$

$$\hat{h}_2 = \tfrac{1}{2} g_{q_1 q_2}^{p_1 p_2} \hat{E}_{q_1 q_2}^{p_1 p_2}, \tag{3c}$$

where h_0 is a constant, and $t_{q_1}^{p_1}$ and $g_{q_1 q_2}^{p_1 p_2}$ are one- and two-electron elements (or integrals), respectively, with all indices summed over. We work in the spin-free form based on the so-called group *generators*, as given by

$$\hat{E}_{q_1}^{p_1} = \sum_{\sigma=\alpha,\beta} \hat{a}_{p_1\sigma}^\dagger \hat{a}_{q_1\sigma}, \tag{4a}$$

$$\hat{E}_{q_1 q_2}^{p_1 p_2} = \sum_{\sigma\tau=\alpha,\beta} \hat{a}_{p_1\sigma}^\dagger \hat{a}_{p_2\tau}^\dagger \hat{a}_{q_2\tau} \hat{a}_{q_1\sigma}, \tag{4b}$$

$$\hat{E}_{q_1 q_2 q_3}^{p_1 p_2 p_3} = \sum_{\sigma\tau\upsilon=\alpha,\beta} \hat{a}_{p_1\sigma}^\dagger \hat{a}_{p_2\tau}^\dagger \hat{a}_{p_3\upsilon}^\dagger \hat{a}_{q_3\upsilon} \hat{a}_{q_2\tau} \hat{a}_{q_1\sigma}, \tag{4c}$$

for the one-, two-, and three-body operators, respectively. Related to these, the reduced density matrices (RDMs) are

introduced as follows,

$$D^{p_1}_{q_1} = \langle \Psi_0 | \hat{E}^{p_1}_{q_1} | \Psi_0 \rangle, \tag{5a}$$

$$D^{p_1 p_2}_{q_1 q_2} = \langle \Psi_0 | \hat{E}^{p_1 p_2}_{q_1 q_2} | \Psi_0 \rangle, \tag{5b}$$

$$D^{p_1 p_2 p_3}_{q_1 q_2 q_3} = \langle \Psi_0 | \hat{E}^{p_1 p_2 p_3}_{q_1 q_2 q_3} | \Psi_0 \rangle, \tag{5c}$$

for the one-, two-, and three-body RDMs, respectively.

Ahead of the transformation eqn (2), the so-called "active-space" description of electron correlation needs be determined as a starting reference electronic state, on top of which dynamic correlation is folded in from the external space when $\hat{\bar{H}}$ in eqn (2) is constructed. We normally employ the CASSCF wavefunction for it, although any type of ansatz can be adopted as long as its reduced density matrices (RDMs) are available. Like other active-space methods, it requires one to model the reference space by dividing an entire set of orbitals (generally indexed by p_i and q_i) into core (doubly-occupied), active (fractionally-occupied), and virtual (unoccupied) orbitals, which are hereafter referred to by the orbital indices c_i, o_i and v_i, respectively. The CASSCF method expands the wavefunction $|\Psi_0\rangle$ into the reference configuration space, covering a full correlation of chemically-relevant active electrons within optimized active orbitals, so as to satisfy the following eigen-equation,

$$\hat{H}^{\mathrm{act}} |\Psi_0\rangle = E_{\mathrm{CASSCF}} |\Psi_0\rangle, \tag{6}$$

through exact diagonalization of active-space Hamiltonian $\hat{H}^{\mathrm{act}} = h^{\mathrm{act}}_0 + \hat{h}^{\mathrm{act}}_1 + \hat{h}^{\mathrm{act}}_2$ where $h^{\mathrm{act}}_0 = h_0 + 2t^{c_1}_{c_1} + (2g^{c_1 c_2}_{c_1 c_2} - g^{c_1 c_2}_{c_2 c_1})$, $\hat{h}^{\mathrm{act}}_1 = \bar{t}^{o_1}_{o_2} \hat{E}^{o_1}_{o_2}$, and $\hat{h}^{\mathrm{act}}_1 = \frac{1}{2} g^{o_1 o_3}_{o_2 o_4} \hat{E}^{o_1 o_3}_{o_2 o_4}$ with $\bar{t}^{o_1}_{o_2} = t^{o_1}_{o_2} + (2g^{o_1 c_3}_{o_2 c_3} - g^{o_1 c_3}_{c_3 o_2})$. Since $|\Psi_0\rangle$ exists only in the active space, E_{CASSCF} is given as an expectation value of not only $\hat{H}^{\mathrm{act}}$ but also $\hat{H}$,

$$\langle \Psi_0 | \hat{H} | \Psi_0 \rangle = E_{\mathrm{CASSCF}} \; (= \langle \Psi_0 | \hat{H}^{\mathrm{act}} | \Psi_0 \rangle), \tag{7}$$

accounting for static correlation energy. In addition, when this reference $|\Psi_0\rangle$ acts on $\hat{\bar{H}}$ [eqn (2)], it yields the CT energy $E_{\mathrm{CASSCF-CT}}$, as follows,

$$\langle \Psi_0 | \hat{\bar{H}} | \Psi_0 \rangle = E_{\mathrm{CASSCF-CT}}, \tag{8}$$

which includes the multireference dynamic correlation energy.

2.2 Amplitude

The singles and doubles CT, termed CT-SD, is our standard model, in which the amplitude operator $\hat{A}$ is written as a sum of one- and two-body operators,

$$\hat{A} = \hat{A}_1 + \hat{A}_2, \tag{9a}$$

$$\begin{cases} \hat{A}_1 = A^{e_1}_{a_1} \hat{\varepsilon}^{e_1}_{a_1} \\ \hat{A}_2 = \frac{1}{2} A^{e_1 e_2}_{a_1 a_2} \hat{\varepsilon}^{e_1 e_2}_{a_1 a_2} \end{cases}, \tag{9b}$$

where anti-Hermitian excitation operators are defined as $\hat{\varepsilon}^{e_1}_{a_1} = (\hat{E}^{e_1}_{a_1} - \hat{E}^{a_1}_{e_1})$ and $\hat{\varepsilon}^{e_1 e_2}_{a_1 a_2} = (\hat{E}^{a_1 a_2}_{e_1 e_2} - \hat{E}^{a_1 a_2}_{e_1 e_2})$, and the indices a_i and e_i run over the joint orbital spaces defined by

$$\{a_i\} = \{c_i\} \oplus \{o_i\}, \tag{10a}$$

$$\{e_i\} = \{o_i\} \oplus \{v_i\}, \tag{10b}$$

respectively. Excitation components of $\hat{A}$ [eqn (9)] are then classified into eleven types, as summarized in Table 1, so that we rewrite it as

$$\begin{aligned} \hat{A} = {} & A^{o_1}_{c_1} \hat{\varepsilon}^{o_1}_{c_1} + A^{o_1 o_2}_{c_1 o_3} \hat{\varepsilon}^{o_1 o_2}_{c_1 o_3} + \frac{1}{2} A^{o_1 o_2}_{c_1 c_2} \hat{\varepsilon}^{o_1 o_2}_{c_1 c_2} + A^{v_1}_{c_1} \hat{\varepsilon}^{v_1}_{c_1} + \frac{1}{2} A^{v_1 v_2}_{c_1 c_2} \hat{\varepsilon}^{v_1 v_2}_{c_1 c_2} \\ & + A^{v_1}_{o_1} \hat{\varepsilon}^{v_1}_{o_1} + A^{o_3 v_1}_{o_1 o_2} \hat{\varepsilon}^{o_3 v_1}_{o_1 o_2} + \frac{1}{2} A^{v_1 v_2}_{o_1 o_2} \hat{\varepsilon}^{v_1 v_2}_{o_1 o_2} \\ & + A^{o_1 v_2}_{c_1 o_2} \hat{\varepsilon}^{o_1 v_2}_{c_1 o_2} + A^{v_1 v_2}_{c_1 o_2} \hat{\varepsilon}^{v_1 v_2}_{c_1 o_2} + A^{o_1 v_2}_{c_1 c_2} \hat{\varepsilon}^{o_1 v_2}_{c_1 c_2}, \end{aligned} \tag{11}$$

in which internal elements $A^{o_2}_{o_1}$ and $A^{o_3 o_4}_{o_1 o_2}$ are not included.

Here we again rewrite eqn (9) (or eqn (11)) in a generalized form as $\hat{A} = \sum_\mu A_\nu \hat{\varepsilon}_\mu$, and then a matrix form of overlap between the operator basis $\{\hat{\varepsilon}_\mu\}$ is given by $S_{\mu\nu} = \langle \hat{\varepsilon}_\mu | \hat{\varepsilon}_\nu \rangle (= \langle \Psi_0 | \hat{\varepsilon}^\dagger_\mu \hat{\varepsilon}_\nu | \Psi_0 \rangle)$. It is readily shown that the matrix $\mathbf{S}$ is not diagonal in multireference setting, and thus the singles and doubles basis that represent $\hat{A}$ are generally non-orthogonal, unlike single-reference formalism. For the sake of numerical stabilization, we alternatively expand $\hat{A}$ in orthogonalized basis operators $\{\hat{\varepsilon}^{\mathrm{orth}}_i\}$,[110] which are given as linear transformation of $\hat{\varepsilon}_\mu$,

$$\hat{\varepsilon}^{\mathrm{orth}}_i = \sum_\mu U_{i\mu} \hat{\varepsilon}_\mu, \tag{12}$$

where $U_{i\mu}$ is the eigenvector (unitary) matrix with which to diagonalize $\mathbf{S}$, *i.e.* $\sum_{\mu\nu} U_{i\mu} U_{j\nu} S_{\mu\nu} = \delta_{ij} \sigma_i$. As seen in Table 2 that shows the elements of $\mathbf{S}$, it is block-diagonal, so that the largest dimension of the block matrices each to be diagonalized is o^3, where o is the number of active orbitals. To remove linear dependencies of the basis, the eigen-components with $\sigma_i < \tau$ (threshold) are truncated. Then, $\hat{A}$ is expanded into orthogonal $\hat{\varepsilon}^{\mathrm{orth}}_i$,

$$\hat{A} = \sum_i A^{\mathrm{orth}}_i \hat{\varepsilon}^{\mathrm{orth}}_i \tag{13}$$

which is equated to eqn (9) *via* $A^{\mathrm{orth}}_i = U_{i\mu} A_\mu$.

Table 1 All types of excitation components for the CTSD amplitude $\hat{A}$ [eqn (11)]

Type	
Core-active	
co	$\hat{\varepsilon}^{o_1}_{c_1} = \hat{E}^{o_1}_{c_1} - \hat{E}^{c_1}_{o_1}$
cooo	$\hat{\varepsilon}^{o_1 o_2}_{c_1 o_3} = \hat{E}^{o_1 o_2}_{c_1 o_3} - \hat{E}^{c_1 o_3}_{o_1 o_2}$
ccoo	$\hat{\varepsilon}^{o_1 o_2}_{c_1 c_2} = \hat{E}^{o_1 o_2}_{c_1 c_2} - \hat{E}^{c_1 c_2}_{o_1 o_2}$
Core-virtual	
cv	$\hat{\varepsilon}^{v_2}_{c_1} = \hat{E}^{v_1}_{c_1} - \hat{E}^{c_1}_{v_1}$
ccvv	$\hat{\varepsilon}^{v_1 v_2}_{c_1 c_2} = \hat{E}^{v_1 v_2}_{c_1 c_2} - \hat{E}^{c_1 c_2}_{v_1 v_2}$
Active-virtual	
ov	$\hat{\varepsilon}^{v_1}_{o_1} = \hat{E}^{v_1}_{o_1} - \hat{E}^{o_1}_{v_1}$
ooov	$\hat{\varepsilon}^{o_3 v_2}_{o_1 o_2} = \hat{E}^{o_3 v_2}_{o_1 o_2} - \hat{E}^{o_1 o_2}_{o_3 v_2}$
oovv	$\hat{\varepsilon}^{v_1 v_2}_{o_1 o_2} = \hat{E}^{v_1 v_2}_{o_1 o_2} - \hat{E}^{o_1 o_2}_{v_1 v_2}$
Core-active-virtual	
ccov	$\hat{\varepsilon}^{o_1 v_2}_{c_1 c_2} = \hat{E}^{o_1 v_2}_{c_1 c_2} - \hat{E}^{c_1 c_2}_{o_1 v_2}$
coov	$\hat{\varepsilon}^{o_1 v_2}_{c_1 o_2} = \hat{E}^{o_1 v_2}_{c_1 o_2} - \hat{E}^{c_1 o_2}_{o_1 v_2}$
covv	$\hat{\varepsilon}^{v_1 v_2}_{c_1 o_2} = \hat{E}^{v_1 v_2}_{c_1 o_2} - \hat{E}^{c_1 o_2}_{v_1 v_2}$

Table 2 Expressions of the overlap matrices between the excitation operator basis $\langle\hat{\varepsilon}_\mu|\hat{\varepsilon}_\nu\rangle$

Type	Overlap
co-co	$\langle\hat{\varepsilon}^{o_1}_{c_1}\|\hat{\varepsilon}^{o'_1}_{c'_1}\rangle = \delta^{c_1}_{c'_1}(2\delta^{o'_1}_{o_1} - D^{o'_1}_{o_1})$
co-cooo	$\langle\hat{\varepsilon}^{o_1}_{c_1}\|\hat{\varepsilon}^{o'_1 o'_2}_{c'_1 o'_3}\rangle = \delta^{c_1}_{c'_1}(2\delta^{o'_1}_{o_1} D^{o'_2}_{o'_3} - \delta^{o'_2}_{o_1} D^{o'_1}_{o'_3} - D^{o'_1 o'_2}_{o_1 o'_3})$
cooo-cooo	$\langle\hat{\varepsilon}^{o_1 o_3}_{c_1 o_3}\|\hat{\varepsilon}^{o'_1 o'_2}_{c'_1 o'_3}\rangle = \delta^{c_1}_{c'_1}\{(2\delta^{o'_1}_{o_1}\delta^{o'_2}_{o_2} - \delta^{o'_1}_{o_2}\delta^{o'_2}_{o_1})D^{o_3}_{o'_3} - D^{o'_1 o_3 o'_2}_{o_1 o_2 o'_3} + 2\delta^{o'_1}_{o_1} D^{o_3 o'_2}_{o_2 o'_3} - \delta^{o'_2}_{o_2} D^{o'_1 o_3}_{o_1 o'_3} - \delta^{o'_2}_{o_1} D^{o'_1 o_3}_{o'_3 o_2} - \delta^{o'_1}_{o_2} D^{o_3 o'_2}_{o_1 o'_3}\}$
ccoo-ccoo	$\langle\hat{\varepsilon}^{o_1 o_2}_{c_1 c_2}\|\hat{\varepsilon}^{o'_1 o'_2}_{c'_1 c'_2}\rangle = \delta^{c_1}_{c'_1}\delta^{c_2}_{c'_2}\{4\delta^{o'_1}_{o_1}\delta^{o'_2}_{o_2} - 2\delta^{o'_1}_{o_2}\delta^{o'_2}_{o_1} + D^{o'_1 o'_2}_{o_1 o_2} - 2(\delta^{o'_1}_{o_1} D^{o'_2}_{o_2} + \delta^{o'_2}_{o_2} D^{o'_1}_{o_1}) + (\delta^{o'_1}_{o_2} D^{o'_2}_{o_1} + \delta^{o'_2}_{o_1} D^{o'_1}_{o_2})\} + [c'_1 \Leftrightarrow c'_2, o'_1 \Leftrightarrow o'_2]$
cv-cv	$\langle\hat{\varepsilon}^{v_1}_{c_1}\|\hat{\varepsilon}^{v'_1}_{c'_1}\rangle = 2\delta^{c_1}_{c'_1}\delta^{v'_1}_{v_1}$
cv-coov	$\langle\hat{\varepsilon}^{v_1}_{c_1}\|\hat{\varepsilon}^{o'_1 v'_2}_{c'_1 o'_1}\rangle = -\delta^{c_1}_{c'_1}\delta^{v'_2}_{v_1} D^{o'_1}_{o'_2}$
cv-covo	$\langle\hat{\varepsilon}^{v_1}_{c_1}\|\hat{\varepsilon}^{v'_2 o'_1}_{c'_1 o'_2}\rangle = 2\delta^{c_1}_{c'_1}\delta^{v'_2}_{v_1} D^{o'_1}_{o'_2}$
coov-coov	$\langle\hat{\varepsilon}^{o_1 v_2}_{c_1 o_2}\|\hat{\varepsilon}^{o'_1 v'_2}_{c'_1 o'_2}\rangle = \delta^{c_1}_{c'_1}\delta^{v_2}_{v'_2}(2\delta^{o'_1}_{o_1} D^{o_2}_{o'_2} - D^{o'_1 o_2}_{o_1 o'_2})$
coov-covo	$\langle\hat{\varepsilon}^{o_1 v_2}_{c_1 o_2}\|\hat{\varepsilon}^{v'_2 o'_1}_{c'_1 o'_2}\rangle = -\delta^{c_1}_{c'_1}\delta^{v_2}_{v'_2}(\delta^{o'_1}_{o_1} D^{o_2}_{o'_2} + D^{o_2 o'_1}_{o_1 o'_2})$
covo-coov	$\langle\hat{\varepsilon}^{v_2 o_1}_{c_1 o_2}\|\hat{\varepsilon}^{o'_1 v'_2}_{c'_1 o'_2}\rangle = -\delta^{c_1}_{c'_1}\delta^{v_2}_{v'_2}(\delta^{o'_1}_{o_1} D^{o_2}_{o'_2} + D^{o_2 o'_1}_{o_1 o'_2})$
covo-covo	$\langle\hat{\varepsilon}^{v_2 o_1}_{c_1 o_2}\|\hat{\varepsilon}^{v'_2 o'_1}_{c'_1 o'_2}\rangle = 2\delta^{c_1}_{c'_1}\delta^{v_2}_{v'_2}(\delta^{o'_1}_{o_1} D^{o_2}_{o'_2} + D^{o_2 o'_1}_{o_1 o'_2})$
ccov-ccov	$\langle\hat{\varepsilon}^{o_1 v_2}_{c_1 c_2}\|\hat{\varepsilon}^{o'_1 v'_2}_{c'_1 c'_2}\rangle = \delta^{v_2}_{v'_2} D^{o'_1}_{o_1}(2\delta^{c_1}_{c'_1}\delta^{c_2}_{c'_2} - \delta^{c_1}_{c'_2}\delta^{c_2}_{c'_1})$
ccvv-ccvv	$\langle\hat{\varepsilon}^{v_1 v_2}_{c_1 c_2}\|\hat{\varepsilon}^{v'_1 v'_2}_{c'_1 c'_2}\rangle = \delta^{v'_1}_{v_1}\delta^{v'_2}_{v_2}(2\delta^{c_1}_{c'_1}\delta^{c_2}_{c'_2} - \delta^{c_1}_{c'_2}\delta^{c_2}_{c'_1}) + [c'_1 \Leftrightarrow c'_2, v'_1 \Leftrightarrow v'_2]$
covv-covv	$\langle\hat{\varepsilon}^{v_1 v_2}_{c_1 o_2}\|\hat{\varepsilon}^{v'_1 v'_2}_{c'_1 o'_2}\rangle = \delta^{c_1}_{c'_1} D^{o_2}_{o'_2}(2\delta^{v'_1}_{v_1}\delta^{v'_2}_{v_2} - \delta^{v'_1}_{v_2}\delta^{v'_2}_{v_1})$
ov-ov	$\langle\hat{\varepsilon}^{v_1}_{o_1}\|\hat{\varepsilon}^{v'_1}_{o'_1}\rangle = \delta^{v'_1}_{v_1} D^{o_1}_{o'_1}$
ov-ooov	$\langle\hat{\varepsilon}^{v_1}_{o_1}\|\hat{\varepsilon}^{o'_3 v'_2}_{o'_1 o'_2}\rangle = \delta^{v'_1}_{v_1} D^{o'_3 o_1}_{o'_1 o'_2}$
ooov-ooov	$\langle\hat{\varepsilon}^{o_3 v_2}_{o_1 o_2}\|\hat{\varepsilon}^{o'_3 v'_2}_{o'_1 o'_2}\rangle = \delta^{v'_1}_{v_1}(\delta^{o'_3}_{o_3} D^{o_1 o_2}_{o'_1 o'_2} - D^{o_1 o_2 o'_3}_{o_3 o'_2 o'_1})$
oovv-oovv	$\langle\hat{\varepsilon}^{v_1 v_2}_{o_1 o_2}\|\hat{\varepsilon}^{v'_1 v'_2}_{o'_1 o'_2}\rangle = \delta^{v'_1}_{v_1}\delta^{v'_2}_{v_2} D^{o_1 o_2}_{o'_1 o'_2} + [o'_1 \Leftrightarrow o'_2, v'_1 \Leftrightarrow v'_2]$

Particular care is taken to the orthogonalization between one-body and semi-internal two-body operators, such as $\hat{\varepsilon}^{v_1}_{o_1}$ and $\hat{\varepsilon}^{o_3 v_2}_{o_1 o_2}$, respectively, which may have nonzero overlap. In some approximation made in CT as mentioned later, one-body orbital rotation arising from semi-internal $\hat{\varepsilon}^{o_3 v_2}_{o_1 o_2}$ cannot be evaluated exactly, whereas it is done exactly with $\hat{\varepsilon}^{v_1}_{o_1}$, explicit one-body operator. Therefore, we project out the pure one-body components from semi-internal two-body operators, which are thus modified to the following intermediates,

$$\begin{aligned}\hat{\varepsilon}'^{o_3 v_2}_{o_1 o_2} &= \hat{\varepsilon}^{o_3 v_2}_{o_1 o_2} - \hat{\varepsilon}^{v_4}_{o_4}\langle\hat{\varepsilon}^{v_4}_{o_4}|\hat{\varepsilon}^{v_5}_{o_5}\rangle^{-1}\langle\hat{\varepsilon}^{v_5}_{o_5}|\hat{\varepsilon}^{o_3 v_2}_{o_1 o_2}\rangle \\ &= \hat{\varepsilon}^{o_3 v_2}_{o_1 o_2} - \hat{\varepsilon}^{v_2}_{o_4}[D^{-1}]^{o_4}_{o_5} D^{o_3 o_5}_{o_1 o_2}.\end{aligned} \tag{14}$$

They fulfill $\langle\hat{\varepsilon}'^{o_3 v_2}_{o_1 o_2}|\hat{\varepsilon}^{v_4}_{o_4}\rangle = 0$, and are orthogonalized by diagonalizing their overlap matrix, leading to a proper style of orthogonal basis.

In the earlier work, we proposed another choice for stabilized basis operators, employing the strong contraction (SC) scheme,[108] which was first introduced by Malrieu *et al.* in the context of NEVPT2 theory[27–30] that addresses instability problems associated with intruder states in CASPT2. In the basis of the SC scheme, the excitation operators $\hat{\varepsilon}_\mu$ are strongly contracted in such a way that SC operators consist of a drastic simplification of the first order interaction basis in which each external orbital (for singles) or orbital pair (for doubles) has only one excitation operator that connects it to the active space. An immediate consequence of this formulation is that SC operators are mutually orthogonal. Importantly, they avoid completely the difficulties in building and diagonalizing overlap matrices and therefore require neither o^9 cost diagonalization step nor the reference function's three-body RDM [eqn (5c)]. Each SC operator is formed as the sum of all contracted operators of its type (*e.g.* double excitations from active space into virtual orbitals v_1 and v_2) weighted by their coefficients in electronic Hamiltonian $\hat{H}$. For the example of double excitations between active and virtual orbitals, the operator corresponding to the pair of virtual orbitals (v_1, v_2) is

$$\hat{\varepsilon}^{\mathrm{SC}}_{v_1 v_2} = \sum_{a_1 a_2} g^{v_1 v_2}_{a_1 a_2}\hat{\varepsilon}^{v_1 v_2}_{a_1 a_2}. \tag{15}$$

The details of the use of strongly contracted excitation operators in CT are described in ref. 108. In this paper, unless otherwise mentioned, we use overlap orthogonalization operators $\hat{\varepsilon}^{\mathrm{orth}}_i$ [eqn (12)] for the amplitude space.

2.3 Operator decomposition

In the CT-SD model, the effective Hamiltonian [eqn (2)] is approximately evaluated in the Baker–Campbell–Hausdorff (BCH) expansion,

$$\begin{aligned}\hat{\bar{H}} &= \hat{H} + [\hat{H},\hat{A}] + \frac{1}{2!}[[\hat{H},\hat{A}],\hat{A}] + \ldots \\ &\approx \hat{H} + [\hat{H},\hat{A}]_{1,2} + \frac{1}{2!}[[\hat{H},\hat{A}]_{1,2},\hat{A}]_{1,2} + \ldots,\end{aligned} \tag{16}$$

where the notation $[\hat{H},\hat{A}]_{1,2}$ indicates that we replace three-body operators resulting from the commutator $[\hat{H},\hat{A}]$ by decompositions into one- and two-body interactions in a way of effectively averaging higher-particle-rank correlation. This approximation, which is recursively applied to the nesting of commutators, allows the infinite BCH expansion of $\hat{\bar{H}}$ to be closed and represented with a linear combination of only one- and two-body operators, and its exponential complexity to be reduced to polynomial computational cost.

The operator decomposition mentioned above is based on Mukherjee and Kutzelnigg's formalism of extended normal ordering (ENO).[115–117] The spin-free form of the three-body operator, $\hat{E}^{p_1 p_2 p_3}_{q_1 q_2 q_3}$ [eqn (4c)], is expressed using the corresponding

ENO, $\hat{\tilde{E}}^{p_1p_2p_3}_{q_1q_2q_3}$, as follows,

$$\begin{aligned}\hat{E}^{p_1p_2p_3}_{q_1q_2q_3} &= \hat{\tilde{E}}^{p_1p_2p_3}_{q_1q_2q_3} + c^{p_1p_2p_3}_{q_1q_2q_3} \\ &+ \{D^{p_1}_{q_1}\mathring{D}^{p_2p_3}_{q_2q_3} - \tfrac{1}{2}D^{p_1}_{q_2}\mathring{D}^{p_2p_3}_{q_1q_3} - \tfrac{1}{2}D^{p_1}_{q_3}\mathring{D}^{p_2p_3}_{q_2q_1} \\ &+ D^{p_2}_{q_2}\mathring{D}^{p_3p_1}_{q_3q_1} - \tfrac{1}{2}D^{p_2}_{q_3}\mathring{D}^{p_3p_1}_{q_2q_1} - \tfrac{1}{2}D^{p_2}_{q_1}\mathring{D}^{p_3p_1}_{q_3q_2} \\ &+ D^{p_3}_{q_3}\mathring{D}^{p_1p_2}_{q_1q_2} - \tfrac{1}{2}D^{p_3}_{q_1}\mathring{D}^{p_1p_2}_{q_3q_2} - \tfrac{1}{2}D^{p_3}_{q_2}\mathring{D}^{p_1p_2}_{q_1q_3}\} \\ &- \{\hat{E}^{p_1}_{q_1}\bar{D}^{p_2p_3}_{q_2q_3} - \tfrac{1}{2}\hat{E}^{p_1}_{q_2}\bar{D}^{p_2p_3}_{q_1q_3} - \tfrac{1}{2}\hat{E}^{p_1}_{q_3}\bar{D}^{p_2p_3}_{q_2q_1} \\ &+ \hat{E}^{p_2}_{q_2}\bar{D}^{p_3p_1}_{q_3q_1} - \tfrac{1}{2}\hat{E}^{p_2}_{q_3}\bar{D}^{p_3p_1}_{q_2q_1} - \tfrac{1}{2}\hat{E}^{p_2}_{q_1}\bar{D}^{p_3p_1}_{q_3q_2} \\ &+ \hat{E}^{p_3}_{q_3}\bar{D}^{p_1p_2}_{q_1q_2} - \tfrac{1}{2}\hat{E}^{p_3}_{q_1}\bar{D}^{p_1p_2}_{q_3q_2} - \tfrac{1}{2}\hat{E}^{p_3}_{q_2}\bar{D}^{p_1p_2}_{q_1q_3}\} \\ &+ \{D^{p_1}_{q_1}\hat{E}^{p_2p_3}_{q_2q_3} - \tfrac{1}{2}D^{p_1}_{q_2}\hat{E}^{p_2p_3}_{q_1q_3} - \tfrac{1}{2}D^{p_1}_{q_3}\hat{E}^{p_2p_3}_{q_2q_1} \\ &+ D^{p_2}_{q_2}\hat{E}^{p_3p_1}_{q_3q_1} - \tfrac{1}{2}D^{p_2}_{q_3}\hat{E}^{p_3p_1}_{q_2q_1} - \tfrac{1}{2}D^{p_2}_{q_1}\hat{E}^{p_3p_1}_{q_3q_2} \\ &+ D^{p_3}_{q_3}\hat{E}^{p_1p_2}_{q_1q_2} - \tfrac{1}{2}D^{p_3}_{q_1}\hat{E}^{p_1p_2}_{q_3q_2} - \tfrac{1}{2}D^{p_3}_{q_2}\hat{E}^{p_1p_2}_{q_1q_3}\}\end{aligned} \tag{17}$$

with

$$\mathring{D}^{p_1p_2}_{q_1q_2} = -D^{p_1p_2}_{q_1q_2} + \tfrac{4}{3}(D^{p_1}_{q_1}D^{p_2}_{q_2} - \tfrac{1}{2}D^{p_1}_{q_2}D^{p_2}_{q_1}), \tag{18}$$

and

$$\bar{D}^{p_1p_2}_{q_1q_2} = -D^{p_1p_2}_{q_1q_2} + 2(D^{p_1}_{q_1}D^{p_2}_{q_2} - \tfrac{1}{2}D^{p_1}_{q_2}D^{p_2}_{q_1}), \tag{19}$$

where $c^{p_1p_2p_3}_{q_1q_2q_3}$ is the three-body cumulant, which physically represents *connected* three-body fluctuation from average one- and two-particle interactions. Note that $\left\langle \hat{\tilde{E}}^{p_1p_2p_3}_{q_1q_2q_3} \right\rangle = 0$ by definition.

Following the cumulant decomposition of RDMs, we can write the three-body RDM element $D^{p_1p_2p_3}_{q_1q_2q_3}$ [eqn (5c)] in terms of products of lower-body RDMs along with the three-body cumulant $c^{p_1p_2p_3}_{q_1q_2q_3}$ as,

$$\begin{aligned}D^{p_1p_2p_3}_{q_1q_2q_3} &= c^{p_1p_2p_3}_{q_1q_2q_3} \\ &+ D^{p_1}_{q_1}\breve{D}^{p_2p_3}_{q_2q_3} - \tfrac{1}{2}D^{p_1}_{q_2}\breve{D}^{p_2p_3}_{q_1q_3} - \tfrac{1}{2}D^{p_1}_{q_3}\breve{D}^{p_2p_3}_{q_2q_1} \\ &+ D^{p_2}_{q_2}\breve{D}^{p_3p_1}_{q_3q_1} - \tfrac{1}{2}D^{p_2}_{q_3}\breve{D}^{p_3p_1}_{q_2q_1} - \tfrac{1}{2}D^{p_2}_{q_1}\breve{D}^{p_3p_1}_{q_3q_2} \\ &+ D^{p_3}_{q_3}\breve{D}^{p_1p_2}_{q_1q_2} - \tfrac{1}{2}D^{p_3}_{q_1}\breve{D}^{p_1p_2}_{q_3q_2} - \tfrac{1}{2}D^{p_3}_{q_2}\breve{D}^{p_1p_2}_{q_1q_3},\end{aligned} \tag{20}$$

where

$$\breve{D}^{p_1p_2}_{q_1q_2} = D^{p_1p_2}_{q_1q_2} - \tfrac{2}{3}(D^{p_1}_{q_1}D^{p_2}_{q_2} - \tfrac{1}{2}D^{p_1}_{q_2}D^{p_2}_{q_1}). \tag{21}$$

The lower-body decomposition ($[\ldots]_{1,2}$), which is the central approximations in the CT theory, is achieved by neglecting the three-body ENO and cumulant, as follows,

$$\hat{\tilde{E}}^{p_1p_2p_3}_{q_1q_2q_3} \Rightarrow 0, \tag{22a}$$

$$c^{p_1p_2p_3}_{q_1q_2q_3} \Rightarrow 0, \tag{22b}$$

for eqn (17) and eqn (20). According to the above formulae for our decomposition, it can be readily shown that the approximate commutator $[\hat{H},\hat{A}]_{1,2}$ takes the following general form:

$$[\hat{H},\hat{A}]_{1,2} = C_0 + \hat{C}_1 + \hat{C}_2, \tag{23a}$$

$$\hat{C}_1 = C_1{}^{p_1}_{q_1}\hat{E}^{p_1}_{q_1}, \tag{23b}$$

$$\hat{C}_2 = \tfrac{1}{2}C_2{}^{p_1p_2}_{q_1q_2}\hat{E}^{p_1p_2}_{q_1q_2}, \tag{23c}$$

and it is, by design, of exactly the same form as the Hamiltonian $\hat{H}$, given by eqn (3).

2.4 Shifted amplitude equation

The amplitudes of $\hat{A}$ (*e.g.* eqn (13)) are determined by solving the following non-linear projected equations, using the iterative Newton–Raphson (NR) method with the typical initial guess $\hat{A} = 0$,

$$R_i = \langle [\hat{\tilde{H}}, \hat{\varepsilon}^{\text{orth}}_i]_{1,2} \rangle = 0, \tag{24}$$

which are regarded as a stationary condition,[109] analogous to its counterparts in coupled-cluster theory. These take the form of the generalized Brillouin conditions.[118]

In this study, as a robust way of avoiding nontrivial convergence issues that arise when solving them, let us introduce an alternative stationary condition, which is formulated as the shifted amplitude equations,

$$R'_i = R_i + \lambda A^{\text{orth}}_i = 0, \tag{25}$$

where λ is a level-shifting parameter. In each NR iteration, the following first-order linear equations are solved to update $\hat{A}$ with the NR step $\Delta\hat{A}$ (*i.e.* $\hat{A} \leftarrow \hat{A} + \Delta\hat{A}$),

$$\sum_j J'_{ij}\Delta A^{\text{orth}}_j = -R'_i, \tag{26}$$

where J'_{ij} is the shifted CT Jacobian matrix, defined by

$$J'_{ij} = J_{ij} + \lambda\delta_{ij}, \tag{27}$$

$$J_{ij} = \langle [[\hat{\tilde{H}}, \hat{\varepsilon}^{\text{orth}}_j]_{1,2}, \hat{\varepsilon}^{\text{orth}}_i]_{1,2} \rangle, \tag{28}$$

with Kronecker delta δ_{ij}.

The diagonal shifting λ terms significantly remedy the condition of the Jacobian matrix, which can have spuriously small eigenvalues of a non-physical nature associated with the operator decomposition, while a constant λ is chosen *ad hoc*. It is closely related to the level shift often employed in multi-reference perturbation theory calculations to regularize singularity associated with intruder states.[119–124]

Although the level shift λ stabilizes the stationary condition, the solution (amplitudes and energy) has to depend on λ, which is of arbitrary choice. Let us consider how we can recover the unshifted solution from the shifted results. As $\Delta\hat{C}$ is defined as the correction to the shifted amplitudes $\hat{A}$, the unshifted solution satisfies the following stationary equation (as originally given in eqn (24))

$$\langle [\mathrm{e}^{-(\hat{A}+\Delta\hat{C})}\hat{H}\mathrm{e}^{(\hat{A}+\Delta\hat{C})}, \hat{\varepsilon}^{\text{orth}}_i]_{1,2} \rangle = 0, \tag{29}$$

Using eqn (25), we arrive at the following equation,

$$\lambda A_i = \sum_j G_{ij}\Delta C_j + O(\Delta C^2), \tag{30}$$

where $G_{ij} = \langle [[H,\hat{\varepsilon}_j] + \tfrac{1}{2}[[H,A],\hat{\varepsilon}_j] + \tfrac{1}{2}[[H,\hat{\varepsilon}_j],A] + \cdots,\hat{\varepsilon}_i] \rangle$. The approximate first-order correction is thus formulated as

$\Delta C \approx \lambda(G + \lambda)^{-1} A + \lambda^2(G + \lambda)^{-2} A + \cdots$, which exhibits a similar structure to the formulas developed in earlier multireference perturbation studies.[119–124] In contrast to multireference perturbation methods, the evaluation of our formula adapted to CT, however, does not seem to be trivial.

3 Parallelized implementation

The computational tasks executed in CT calculations can be broken down into roughly two types of repeatedly-called subroutines, namely, those to evaluate (1) commutator $[\hat{H},\hat{A}]_{1,2}$ and (2) residual elements $\langle[\hat{H},\hat{\varepsilon}_i]_{1,2}\rangle$. They are the most computationally intensive and thus pose a challenge to implementation. We approach this using a parallel algorithm which distributes the floating point operations and some of the storage across multiple computer processors.

Our central parallelization strategy designed for the distributed architecture, which is a prevailing parallel computing platform, is to distribute the storage of four-index arrays of:

1. General two-body elements (called 'H2 tensor'), *e.g.* $g^{p_1p_2}_{q_1q_2}$ and $C_2{}^{p_1p_2}_{q_1q_2}$ for Hamiltonian $\hat{H}$ [eqn (3)] and the commutator $[\hat{H},\hat{A}]_{1,2}$ [eqn (23)], respectively, and

2. Two-body amplitude elements (called 'A2 tensor'), *e.g.* $A^{e_1e_2}_{a_1a_2}$ and $R^{e_1e_2}_{a_1a_2}$ for the amplitude $\hat{A}$ [eqn (9)] and the residual elements $\langle[\hat{H},\hat{\varepsilon}^{e_1e_2}_{a_1a_2}]_{1,2}\rangle$ [eqn (24)], respectively.

This strategy is motivated by the n^4 and $n^2_{\rm act}n^2_{\rm ext}$ sizes for the H2 and A2 tensors, respectively, where n is the number of total orbitals, indexed by p_i and q_i, and $n_{\rm act}$ is the number of core and active orbitals, indexed by a_i [eqn (10a)], while $n_{\rm ext}$ is the number of active and virtual orbitals, indexed by e_i [eqn (10b)]. They rapidly become too large to store in the fast memory of a single processor.

For the H2 tensors (*e.g.* $g^{p_1p_2}_{q_1q_2}$), we have implemented such storage distribution in a way of splitting each tensor by its p_1 index, so that each processor stores the tensor elements for a limited set of values of p_1 and all values of p_2, q_1, and q_2. To further reduce memory requirements, we store the chopped tensor on the hard disk of each assigned owner processor, loading its elements into fast memory one n^3 sized block at a time. A block is defined as the set of all two-body tensor elements associated with a specific value of p_1, say, it can correspond to the FORTRAN array slicing g $(1:n, 1:n, 1:n,p_1)$. As shown later, this data-parallel model for the H2 tensor allows us to evenly distribute CT theory's floating point operations across our processors.

The A2 tensors (*e.g.* $A^{e_1e_2}_{a_1a_2}$) are also physically distributed across processors by splitting each tensor by its e_2 index running over the sub-range. This parallelism has been implemented by using the Global Arrays (GA) toolkit[125] that provides a shared memory style programming environment in the context of distributed array data structures. It facilitates coding of the distributed data algorithm in which each process asynchronously accesses *remote* data blocks of the A2 tensors *via get*, *put*, and *accumulate* operations with "one-sided"-type communications underneath. With this toolkit, the required size of fast memory to store a single A2 tensor results in approximately $(n^2_{\rm act}n^2_{\rm ext}/N_{\rm proc})$ words per processor. Our code is further able to exploit molecular point group symmetry to reduce the memory usage of the chopped A2 tensor, in which the elements are zero and thus unallocated in memory unless the product of the irreducible representations of their a_1, a_2, e_1, and e_2 indices is totally symmetric.

Table 3 Distribution and storage of Hamiltonian $\hat{H}$, commutator $[\hat{H},\hat{A}]_{1,2}$, amplitudes $\hat{A}$, residuals $\hat{R}$, and reduced density matrix elements

Variables		Memory size per processor
$\hat{H}$ [eqn (3)] and $[\hat{H},\hat{A}]_{1,2}$ [eqn (23)]		
h_0, C_0	Replicated	1
$t^{p_1}_{q_1}$, $C_1{}^{p_1}_{q_1}$	Replicated	n^2
$g^{p_1p_2}_{q_1q_2}$, $C_2{}^{p_1p_2}_{q_1q_2}$	Distributed	n^3 (for accessing a block of $n^4/N_{\rm proc}$ elements on disk)
$\hat{A}$ [eqn (9)] and $\hat{H},\hat{\varepsilon}_i]_{1,2}\rangle$ [eqn (24)]		
$A^{e_1}_{a_1}$, $R^{e_1}_{a_1}$	Replicated	$n_{\rm act}n_{\rm ext}$
$A^{e_1e_2}_{a_1a_2}$, $R^{e_1e_2}_{a_1a_2}$	Distributed[a]	$n^2_{\rm act}\,n^2_{\rm ext}/N_{\rm proc}$
Density matrices [eqn (5a) and (5b)]		
$D^{o_1}_{o_2}$	Replicated	$n^2_{\rm occ}$
$D^{o_1o_2}_{o_3o_4}$	Replicated	$n^4_{\rm occ}$

[a] Data-parallel allocation using the global arrays routine. The memory size is further reduced where the molecular symmetry is available.

The remaining tensors (h_0, C_0, $t^{p_1}_{q_1}$, $C_1{}^{p_1}_{q_1}$, $D^{o_1}_{o_2}$, $D^{o_1o_2}_{o_3o_4}$, $A^{e_1}_{a_1}$, $R^{e_1}_{a_1}$) are stored redundantly in the fast memory of each processor as replicated data. Table 3 shows a summary regarding the storage allocations for these tensors.

In what follows, we describe the details of our implementation to evaluate in parallel the commutator $[\hat{H},\hat{A}]_{1,2}$ [eqn (23)] and residual elements $\langle[\hat{H},\hat{\varepsilon}_i]_{1,2}\rangle$ [eqn (24)], along with the explicit tensor contraction formulae.

3.1 Commutators

Let us write the approximate commutator eqn (23) by separating $\hat{H}$ and $\hat{A}$ into their one- and two-body components as,

$$[\hat{H},\hat{A}]_{1,2} = [\hat{h}_1,\hat{A}_1]_{1,2} + [\hat{h}_1,\hat{A}_2]_{1,2} + [\hat{h}_2,\hat{A}_1]_{1,2} + [\hat{h}_2,\hat{A}_2]_{1,2}. \quad (31)$$

The coefficients C_0, $C_1{}^{p_1}_{q_1}$, and $C_2{}^{p_1p_2}_{q_1q_2}$, introduced in eqn (23), are each decomposed into the contributions from these four components as follows,

$$C_0 = c_0, \quad (32a)$$

$$C_1{}^{p_1}_{q_1} = c_1{}^{p_1}_{q_1} + c'_1{}^{p_1}_{q_1}, \quad (32b)$$

$$C_2{}^{p_1p_2}_{q_1q_2} = c_2{}^{p_1p_2}_{q_1q_2} + c'_2{}^{p_1p_2}_{q_1q_2}, + c''_2{}^{p_1p_2}_{q_1q_2}. \quad (32c)$$

The expressions to evaluate the above decomposed elements, c_0, c_1, c'_1, c_2, c'_2, and c''_2, are shown below.

We begin by formulating the simplest commutator $[\hat{h}_1,\hat{A}_1]_{1,2}$, which is expressed as,

$$[\hat{h}_1,\hat{A}_1]_{1,2} = t^{p_1}_{q_1}A^{e_1}_{a_1}[\hat{E}^{p_1}_{q_1},\hat{E}^{e_1}_{a_1} - \hat{E}^{a_1}_{e_1}]_{1,2} = c_1{}^{p_1}_{q_1}\hat{E}^{p_1}_{q_1}, \quad (33)$$

where the matrix $c_1{}^{p_1}_{q_1}$ is given by the symmetrization of the matrix $\bar{c}_1{}^{p_1}_{q_1}$,

$$c_1{}^{p_1}_{q_1} = \tfrac{1}{2}(\bar{c}_1{}^{p_1}_{q_1} + \bar{c}_1{}^{q_1}_{p_1}), \quad (34)$$

and the matrix elements $\bar{c}_1{}^{p_1}_{q_1}$ are determined from the following matrix products,

$$\bar{c}_1{}^{p_1}_{q_1} = 2(t^{p_1}_{e_1} A^{e_1}_{a_1})\delta^{q_1}_{a_1} - 2(t^{p_1}_{a_1} A^{e_1}_{a_1})\delta^{q_1}_{e_1}. \tag{35}$$

Next, the commutator $[\hat{h}_1,\hat{A}_2]_{1,2}$ is written as

$$[\hat{h}_1, \hat{A}_2]_{1,2} = \tfrac{1}{2} t^{p_1}_{q_1} A^{e_1 e_2}_{a_1 a_2} [\hat{E}^{p_1}_{q_1}, \hat{E}^{e_1 e_2}_{a_1 a_2} - \hat{E}^{a_1 a_2}_{e_1 e_2}]_{1,2} = \tfrac{1}{2} c_2{}^{p_1 p_2}_{q_1 q_2} \hat{E}^{p_1 p_2}_{q_1 q_2}, \tag{36}$$

where the symmetrization for the four-index array $c_2{}^{p_1 p_2}_{q_1 q_2}$ is carried out as follows,

$$c_2{}^{p_1 p_2}_{q_1 q_2} = \tfrac{1}{4}(\bar{c}_2{}^{p_1 p_2}_{q_1 q_2} + \bar{c}_2{}^{p_2 p_1}_{q_2 q_1} + \bar{c}_2{}^{q_1 q_2}_{p_1 p_2} + c_2{}^{q_2 q_1}_{p_2 p_1}), \tag{37}$$

and the formula to calculate the array elements $\bar{c}_2{}^{p_1 p_2}_{q_1 q_2}$ is given by

$$c_2{}^{p_1 p_2}_{q_1 q_2} = 4(t^{p_1}_{e_1} A^{e_1 e_2}_{a_1 a_2})\delta^{p_2}_{e_2}\delta^{q_1}_{a_1}\delta^{q_2}_{a_2} - 4(t^{p_1}_{a_1} A^{e_1 e_2}_{a_1 a_2})\delta^{p_2}_{a_2}\delta^{q_1}_{e_1}\delta^{q_2}_{e_2}. \tag{38}$$

The parallel computing of the necessary tensor contractions eqn (35) and eqn (38) for $[\hat{h}_1,\hat{A}_1]$ and $[\hat{h}_1,\hat{A}_2]$, respectively, can now be feasibly implemented because, as already described, each processor redundantly owns all tensors in these terms except the H2 tensor $\bar{c}_2{}^{p_1 p_2}_{q_1 q_2}$, which nevertheless can be evaluated independently for each block associated with p_1 at its owner processor.

The commutators involving $\hat{h}_2$ are described as

$$[\hat{h}_2, \hat{A}_1]_{1,2} = \tfrac{1}{2} g^{p_1 p_2}_{q_1 q_2} A^{e_1}_{a_1} [\hat{E}^{p_1 p_2}_{q_1 q_2}, \hat{E}^{e_1}_{a_1} - \hat{E}^{a_1}_{e_1}]_{1,2} = \tfrac{1}{2} c'_2{}^{p_1 p_2}_{q_1 q_2} \hat{E}^{p_1 p_2}_{q_1 q_2}, \tag{39}$$

$$\begin{aligned}[\hat{h}_2, \hat{A}_1]_{1,2} &= \tfrac{1}{2} g^{p_1 p_2}_{q_1 q_2} A^{e_1 e_2}_{a_1 a_2} [\hat{E}^{p_1 p_2}_{q_1 q_2}, \hat{E}^{e_1 e_2}_{a_1 a_2} - \hat{E}^{a_1 a_2}_{e_1 e_2}]_{1,2} \\ &= c_0 + c'_1{}^{p_1}_{q_1} \hat{E}^{p_1}_{q_1} + \tfrac{1}{2} c''_2{}^{p_1 p_2}_{q_1 q_2} \hat{E}^{p_1 p_2}_{q_1 q_2}.\end{aligned} \tag{40}$$

Evaluating these terms is more challenging, as they each have a couple of the four-index H2 tensors for the input and output data arrays, $g^{p_1 p_2}_{q_1 q_2}$ and $c'_2{}^{p_1 p_2}_{q_1 q_2}$ (or $c''_2{}^{p_1 p_2}_{q_1 q_2}$). At first glance, evaluating these terms requires access to all blocks of c' (or c'') for each block of g. However, a careful inspection of these terms and the 4-fold symmetries of g and c' (or c'') ($g^{p_1 p_2}_{p_3 p_4} = g^{p_2 p_1}_{p_4 p_3} = g^{p_3 p_4}_{p_1 p_2} = g^{p_4 p_3}_{p_2 p_1}$) reveals that each block of g can be made to contribute only to the same block of c' (or c''), as long as c' (or c'') is symmetrized afterwards. This means that a single index of each of the four-index arrays $g^{p_1 p_2}_{q_1 q_2}$, $\bar{c}'_2{}^{p_1 p_2}_{q_1 q_2}$, and $\bar{c}''_2{}^{p_1 p_2}_{q_1 q_2}$ can be shared or common. To emphasize this, the common index is denoted as a bold index $\mathbf{p_1}$. We shall thus say that, for a given $\mathbf{p_1}$, the output elements $\bar{c}'_2{}^{\mathbf{p_1} p_2}_{q_1 q_2}$ and $\bar{c}''_2{}^{\mathbf{p_1} p_2}_{q_1 q_2}$ ($\forall\, p_2, q_1, q_2$) are calculated from the input array $g^{\mathbf{p_1} p'_2}_{q'_1 q'_2}(\forall p'_2, q'_1, q'_2)$, and thereby the required storage in fast memory is $O(n^3)$. This is exemplified clearly in the following tensor contraction to evaluate $[\hat{h}_2,\hat{A}_1]_{1,2}$ [eqn (39)],

$$\bar{c}'_2{}^{\mathbf{p_1} p_2}_{q_1 q_2} = 4(v^{\mathbf{p_1} p_2}_{e_1 q_2} A^{e_1}_{a_1})\delta^{q_1}_{a_1} - 4(v^{\mathbf{p_1} p_2}_{a_1 q_2} A^{e_1}_{a_1})\delta^{q_1}_{e_1}, \tag{41}$$

which is to be symmetrized, resulting in $\bar{c}'_2{}^{p_1 p_2}_{q_1 q_2}$. This trick is exploited in our implementation to achieve memory savings and a simple parallelization, which means that each processor can evaluate these terms from the associated block of g without any network communication for accessing the necessary blocks of the H2 tensors. The rough sketch of our implementation to evaluate $\bar{c}'_2{}^{\mathbf{p_1} p_2}_{q_1 q_2}$ elements given by eqn (41) is:

1. The orbital range for p_1 is divided evenly into sub-ranges, with each sub-range assigned to a different processor. Thus the two-body integrals $g^{x_1 p_2}_{p_3 p_4}\ \forall p_2, p_3, p_4$ are stored in the hard disk of the processor assigned x_1.
2. On each processor, select x_1 from the assigned range.
3. Load the n^3 sized block of two-body integrals $g^{x_1 p_2}_{p_3 p_4}\ \forall p_2, p_3, p_4$ into the processor's fast memory.
4. Evaluate the tensor contractions [eqn (41)] to produce the block of $\bar{c}'_2{}^{x_1 q_2}_{q_3 q_4}\ \forall q_2, q_3, q_4$.
5. Write the two-body block $\bar{c}'_2{}^{x_1 q_2}_{q_3 q_4}\ \forall q_2, q_3, q_4$ to the processor's hard disk.
6. Go to step 2 until all values of x_1 belonging to this processor have been exhausted.
7. Symmetrize the H2 tensor $\bar{c}'_2$ across the processors. This step requires an n^4/n_{proc} amount of data transfer.

For $[\hat{h}_2,\hat{A}_2]_{1,2}$ [eqn (40)], the tensor product formulas to evaluate the coefficients c_0, $\bar{c}'_1{}^{p_1}_{q_1}$, $\bar{c}''_2{}^{p_1 p_2}_{q_1 q_2}$ are shown below. Bear in mind that some of them need be symmetrized like eqn (34) and (37) for $\bar{c}_1{}^{p_1}_{q_1}$ and $\bar{c}_2{}^{p_1 p_2}_{q_1 q_2}$, respectively. First, we now have

$$\begin{aligned}c_0 &= 2(g^{p_1 p_2}_{e_1 q_2} A^{e_1 e_2}_{a_1 a_2}) \overset{\circ}{D}{}^{p_1 p_2 e_2}_{a_1 q_2 a_2} \\ &\quad - 2(g^{p_1 p_2}_{a_1 q_2} A^{e_1 e_2}_{a_1 a_2}) \overset{\circ}{D}{}^{p_1 p_2 a_2}_{e_1 q_2 e_2},\end{aligned} \tag{42}$$

where the definition of $\overset{\circ}{D}{}^{p_1 p_2 p_3}_{q_1 q_2 q_3}$ is given from that of $D^{p_1 p_2 p_3}_{q_1 q_2 q_3}$ [eqn (20)] by replacing $\breve{D}^{p_1 p_2}_{q_1 q_2}$ (eqn (21)) in it with $\overset{\circ}{D}{}^{p_1 p_2}_{q_1 q_2}$ [eqn (18)]. We may also evaluate

$$\bar{c}'_1{}^{p_1}_{q_1} = 2(e'_1{}^{p_1}_{q_1} - a'_1{}^{p_1}_{q_1}), \tag{43}$$

in which various useful intermediates are defined as

$$\begin{aligned}e'_1{}^{p_1}_{q_1} &= -\delta^{q_1}_{a_2}(g^{p_1 I}_{e_2 J} S_0{}^{e_2 J}_{a_2 J} + g^{p_1 I}_{J e_2} S_1{}^{e_2 J}_{a_2 J}) \\ &\quad - (g^{p_1 I}_{q_1 e_1} - \tfrac{1}{2} g^{p_1 I}_{e_1 q_1}) S_2{}^{e_1}_{I} + \delta^{q_1}_{e_2}(g^{IJ}_{e_1 p_1} S_4{}^{e_1 e_2}_{IJ}) \\ &\quad - \delta^{p_1}_{a_1}\delta^{q_1}_{e_1}(A^{e_1 e_2}_{a_1 a_2} - \tfrac{1}{2} A^{e_1 e_2}_{a_2 a_1}) S_6{}^{a_2}_{e_2},\end{aligned} \tag{44}$$

and

$$\begin{aligned}a'_1{}^{p_1}_{q_1} &= -\delta^{q_1}_{e_2}(g^{p_1 I}_{a_2 J} S_0{}^{e_2 J}_{a_2 J} + g^{p_1 I}_{J a_2} S_1{}^{e_2 J}_{a_2 J}) \\ &\quad - (g^{p_1 I}_{q_1 a_1} - \tfrac{1}{2} g^{p_1 I}_{a_1 q_1}) S_3{}^{I}_{a_1} \\ &\quad + \delta^{q_1}_{a_2}(g^{IJ}_{a_1 p_1} S_5{}^{IJ}_{a_1 a_2}) \\ &\quad - \delta^{p_1}_{a_1}\delta^{q_1}_{e_1}(A^{e_1 e_2}_{a_1 a_2} - \tfrac{1}{2} A^{e_1 e_2}_{a_2 a_1}) S_6{}^{e_2}_{a_2},\end{aligned} \tag{45}$$

where

$$S_0{}^{e_2 I}_{a_2 J} = [A^{e_1 e_2}_{a_1 a_2} - \tfrac{1}{2} A^{e_1 e_2}_{a_2 a_1}] \bar{D}^{e_1 I}_{a_1 J} \tag{46a}$$

$$S_1{}^{e_2 I}_{a_2 J} = -\tfrac{1}{2}[A^{e_1 e_2}_{a_2 a_1} \bar{D}^{e_1 I}_{J a_1}] - \tfrac{1}{2}[A^{e_1 e_2}_{a_1 a_2} \bar{D}^{e_1 I}_{a_1 J}], \tag{46b}$$

$$S_2{}^{e_1}_{I} = [A^{e_1 e_2}_{a_1 a_2} \bar{D}^{I e_2}_{a_1 a_2}], \tag{46c}$$

$$S_3{}^{I}_{a_1} = [A^{e_1 e_2}_{a_1 a_2} \bar{D}^{I a_2}_{e_1 e_2}], \tag{46d}$$

$$S_4{}_{IJ}^{e_1e_2} = \tfrac{1}{2}[A_{a_1a_2}^{e_1e_2}\bar{D}_{a_1a_2}^{IJ}], \tag{46e}$$

$$S_5{}_{a_1a_2}^{IJ} = \tfrac{1}{2}[A_{a_1a_2}^{e_1e_2}\bar{D}_{e_1e_2}^{IJ}], \tag{46f}$$

$$S_6{}_{p_1}^{a_1} = [g_{q_1q_2}^{p_1p_2}\bar{D}_{q_1q_2}^{a_1p_2}]. \tag{46g}$$

Note that the intermediates S_0, S_1,…, S_6 can be pre-computed, and the size to store them in fast memory is $O(n_{\text{act}}n_{\text{ext}}^2)$.

Finally, we may evaluate

$$\bar{c}''{}_2{}_{q_1q_2}^{\mathbf{p_1}p_2} = 4(e''{}_2{}_{q_1q_2}^{\mathbf{p_1}p_2} - a''{}_2{}_{q_1q_2}^{\mathbf{p_1}p_2}), \tag{47}$$

using the intermediates

$$\begin{aligned} e''{}_2{}_{q_1q_2}^{\mathbf{p_1}p_2} &= \delta_{a_1}^{q_1}\delta_{a_2}^{q_2}(\tfrac{1}{2}g_{e_1e_2}^{\mathbf{p_1}p_2}A_{a_1a_2}^{e_1e_2}) \\ &+ \delta_{a_2}^{p_2}\delta_{e_2}^{q_2}\{g_{q_1I}^{\mathbf{p_1}e_1}(T_0{}_{Ia_2}^{e_1e_2} - \tfrac{1}{2}T_0{}_{Ia_2}^{e_2e_1}) \\ &- \tfrac{1}{2}(g_{Iq_1}^{\mathbf{p_1}e_1}T_0{}_{Ia_2}^{e_1e_2})\} \\ &- \delta_{a_2}^{p_2}\delta_{e_2}^{q_1}(\tfrac{1}{2}g_{Iq_2}^{\mathbf{p_1}e_1}T_0{}_{Ia_2}^{e_2e_1}) \\ &- \delta_{a_2}^{q_1}\delta_{e_2}^{q_2}(\tfrac{1}{2}g_{e_2I}^{\mathbf{p_1}p_2}T_1{}_{a_1a_2}^{Ie_2}) \\ &+ \delta_{a_2}^{q_1}(g_{e_2q_2}^{\mathbf{p_1}p_2}T_2{}_{a_2}^{e_2}) \\ &+ \delta_{a_1}^{q_1}\delta_{a_2}^{q_2}\delta_{e_2}^{p_2}(A_{a_1a_2}^{e_1e_2}T_3{}_{e_1}^{\mathbf{p_1}}), \end{aligned} \tag{48}$$

and

$$\begin{aligned} a''{}_2{}_{q_1q_2}^{\mathbf{p_1}p_2} &= \delta_{e_1}^{q_1}\delta_{e_2}^{q_2}(\tfrac{1}{2}g_{a_1a_2}^{\mathbf{p_1}p_2}A_{a_1a_2}^{e_1e_2}) \\ &+ \delta_{e_2}^{p_2}\delta_{a_2}^{q_2}\{g_{q_1I}^{\mathbf{p_1}a_1}(T_1{}_{a_1a_2}^{Ie_2} - \tfrac{1}{2}T_1{}_{a_2a_1}^{Ie_2}) \\ &- \tfrac{1}{2}(g_{Iq_1}^{\mathbf{p_1}a_1}T_1{}_{a_1a_2}^{Ie_2})\} \\ &- \delta_{e_2}^{p_2}\delta_{a_2}^{q_1}(\tfrac{1}{2}g_{Iq_2}^{\mathbf{p_1}a_1}T_1{}_{a_2a_1}^{Ie_2}) \\ &- \delta_{e_2}^{q_1}\delta_{e_1}^{q_2}(\tfrac{1}{2}g_{a_2I}^{\mathbf{p_1}p_2}T_0{}_{Ia_2}^{e_1e_2}) \\ &+ \delta_{e_2}^{q_1}(g_{a_2q_2}^{\mathbf{p_1}p_2}T_2{}_{a_2}^{e_2}) \\ &+ \delta_{e_1}^{q_1}\delta_{e_2}^{q_2}\delta_{a_2}^{p_2}(A_{a_1a_2}^{e_1e_2}T_3{}_{a_1}^{\mathbf{p_1}}), \end{aligned} \tag{49}$$

where

$$T_0{}_{Ia_2}^{e_1e_2} = [A_{a_1a_2}^{e_1e_2}D_{a_1}^{I}], \tag{50a}$$

$$T_1{}_{a_1a_2}^{Ie_2} = [A_{a_1a_2}^{e_1e_2}D_{e_1}^{I}], \tag{50b}$$

$$T_2{}_{a_2}^{e_2} = [A_{a_1a_2}^{e_1e_2} - \tfrac{1}{2}A_{a_2a_1}^{e_2e_1}]D_{a_1}^{e_1}, \tag{50c}$$

$$T_3{}_{q_1}^{\mathbf{p_1}} = [g_{q_1q_2}^{\mathbf{p_1}p_2} - \tfrac{1}{2}g_{q_2q_1}^{\mathbf{p_1}p_2}]D_{q_2}^{p_2}. \tag{50d}$$

As before, the tensors T_0, …, T_3 can be precomputed and stored in fast memory, whose usage is $O(n_{\text{act}}^2\, n_{\text{ext}})$.

For the programming of these tensor contractions, there are simplifications for the RDMs involving the core orbital indices, $D_{c_2}^{c_1} = 2\delta_{c_2}^{c_1}$, $D_{c_3o_4}^{c_1o_2} = D_{c_3}^{c_1}D_{o_4}^{o_2}$, $D_{o_3c_4}^{c_1o_2} = -\frac{1}{2}D_{c_4}^{c_1}D_{o_3}^{o_2}$, and $D_{c_3c_4}^{c_1c_2} = D_{c_3}^{c_1}D_{c_4}^{c_2} - \frac{1}{2}D_{c_4}^{c_1}D_{c_3}^{c_2}$. The parallel algorithm to evaluate the commutator is summarized in Algorithm 1.

Algorithm 1 Parallel algorithm to evaluate the commutator

```
Evaluate c̄1 p1_q1 [eqn (35)] ∀ p1,q1 for [ĥ1,Â1]1,2.
for p1 ∈ proc do
  Load g^{p1p2}_{q1q2} ∀p2, q1, q2.
  Evaluate c̄'2 p1p2_q1q2 [eqn (41)] ∀ p2,q1,q2 for [ĥ2,Â1]1,2.
  for e2 do
  Load A^{e1e2}_{a1a2} ∀a1, a2, e1 from GA
  Evaluate c̄2 p1p2_q1q2 [eqn (38)] ∀ p2,q1,q2 for [ĥ1,Â2]1,2.
  Evaluate c0 [eqn (42)], c̄'1 p'1_q1 [eqn (43)] ∀p'1, q1, c̄''2 p1p2_q1q2
  [eqn (47)] ∀ p2,q1,q2 for [ĥ2,Â2]1,2.
  end for
  C0 += c0.
  C̄1 p'1_q1 += c̄1 p'1_q1 + c̄'1 p'1_q1 ∀p'1, q1.
  Load C̄^{p1p2}_{q1q2} ∀p2, q1, q2.
  C̄^{p1p2}_{q1q2} += c̄2 p1p2_q1q2 + c̄'2 p1p2_q1q2 + c̄''2 p1p2_q1q2 ∀p2, q1, q2.
  Save C̄^{p1p2}_{q1q2} ∀p2, q1, q2.
end for
Symmetrize C̄1 p1_q1 and C̄^{p1p2}_{q1q2}.
```

3.2 Residual elements

The residual elements are given by

$$R_{a_1}^{e_1} = \langle[\hat{\bar{H}}, \hat{E}_{a_1}^{e_1}]_{1,2}\rangle, \tag{51a}$$

$$R_{a_1a_2}^{e_1e_2} = \langle[\hat{\bar{H}}, \hat{E}_{a_1a_2}^{e_1e_2}]_{1,2}\rangle, \tag{51b}$$

and are decomposed into the contributions from the 1- and 2-body operators of $\hat{\bar{H}}$,

$$R_{a_1}^{e_1} = R'_1{}_{a_1}^{e_1} + R''_1{}_{a_1}^{e_1}, \tag{52a}$$

$$R_{a_1a_2}^{e_1e_2} = R'_2{}_{a_1a_2}^{e_1e_2} + R''_2{}_{a_1a_2}^{e_1e_2}, \tag{52b}$$

where $R'_1{}_{a_1}^{e_1}$ and $R'_2{}_{a_1a_2}^{e_1e_2}$ are associated with the 1-body operator $\hat{h}_1$ [eqn (3b)], and $R''_1{}_{a_1}^{e_1}$ and $R''_2{}_{a_1a_2}^{e_1e_2}$ with the 2-body operator $\hat{h}_2$ [eqn (3c)]. The tensor product forms of $R'_1{}_{a_1}^{e_1}$ and $R'_2{}_{a_1a_2}^{e_1e_2}$ are given by

$$\begin{aligned} R'_1{}_{a_1}^{e_1} &= \langle[\hat{h}_1, \hat{E}_{a_1}^{e_1} - \hat{E}_{e_1}^{a_1}]_{1,2}\rangle \\ &= 2(t_{e_1}^{p_1}D_{a_1}^{p_1} - t_{a_1}^{p_1}D_{e_1}^{p_1}), \end{aligned} \tag{53}$$

$$\begin{aligned} R'_2{}_{a_1a_2}^{e_1e_2} &= \langle[\hat{h}_1, \hat{E}_{a_1a_2}^{e_1e_2} - \hat{E}_{e_1e_2}^{a_1a_2}]_{1,2}\rangle \\ &= \tfrac{1}{2}(\bar{R}'_2{}_{a_1a_2}^{e_1e_2} + \bar{R}'_2{}_{a_2a_1}^{e_2e_1}), \end{aligned} \tag{54}$$

$$R'_2{}_{a_1a_2}^{e_1e_2} = 2(t_{e_1}^{p_1}D_{a_1a_2}^{p_1e_2} - t_{a_1}^{p_1}D_{e_1e_2}^{p_1a_2}), \tag{55}$$

which can be easily implemented as the multiplications between the matrix $t_{q_1}^{p_1}$ [eqn (3b)] and the 1- and 2-body RDMs, which are all kept in fast memory. The expressions to evaluate

$R''^{e_1}_{1\ a_1}$ and $R''^{e_1e_2}_{2\ a_1a_2}$ are given by

$$R''^{e_1}_{1\ a_1} = \langle \Psi_0|[\hat{h}_2, \hat{E}^{e_1}_{a_1} - \hat{E}^{a_1}_{e_1}]_{1,2}|\Psi_0\rangle = 2(v^{p_1p_2}_{e_1q_2} D^{p_1p_2}_{a_1q_2} - v^{p_1p_2}_{a_1q_2} D^{p_1p_2}_{e_1q_2}), \quad (56)$$

$$R''^{e_1e_2}_{2\ a_1a_2} = \langle \Psi_0|[h_2, \hat{E}^{e_1e_2}_{a_1a_2} - \hat{E}^{a_1a_2}_{e_1e_2}]_{1,2}|\Psi_0\rangle = \tfrac{1}{2}(\bar{R}''^{e_1e_2}_{2\ a_1a_2} + \bar{R}''^{e_1e_2}_{2\ a_1a_2}) \quad (57)$$

$$\bar{R}''^{e_1e_2}_{2\ a_1a_2} = (v^{p_1p_2}_{e_1e_2} D^{p_1p_2}_{a_1a_2} - v^{p_1p_2}_{a_1a_2} D^{p_1p_2}_{e_1e_2}) + 2(v^{p_1p_2}_{e_1q_2} D^{p_1p_2e_2}_{a_1q_2a_2} - v^{p_1p_2}_{a_1q_2} D^{p_1p_2a_2}_{e_1q_2e_2}). \quad (58)$$

The last two terms $v^{p_1p_2}_{e_1q_2} D^{p_1p_2e_2}_{a_1q_2a_2}$ and $v^{p_1p_2}_{a_1q_2} D^{p_1p_2a_2}_{e_1q_2e_2}$ are each seven-fold summations which at first glance have an $O(n^7)$ evaluation cost. This expense, however, can be reduced to $O(n^6)$ by using the decomposed form of the 3-body RDM $D^{p_1p_2e_2}_{a_1q_2a_2}$, given by eqn (20).

Once all processors have evaluated their individual contributions, these are summed together. This final summation requires an $n^2_{\mathrm{occ}}\, n^2_{\mathrm{ext}} \log n_{\mathrm{proc}}$ amount of data transfer between processors, where n_{proc} is the number of processors.

4 Numerical results

The performance of the level-shifted condition [eqn (25)], introduced in this work, has been assessed. We performed benchmark CT calculations for the symmetric breaking of the water molecule and the bond breaking curve of the nitrogen molecule. These molecules were chosen for a direct comparison with the results in our previous papers.[109,110] We used 6-31G and cc-pVDZ basis sets for H_2O and N_2, respectively, and CAS(6e,5o) [H_2O] and CAS(6e,6o) [N_2] for the CASSCF references. Fig. 1 and 2 show the errors in the total energies as measured from the FCI (for H_2O) or MRCI+Q (for N_2)

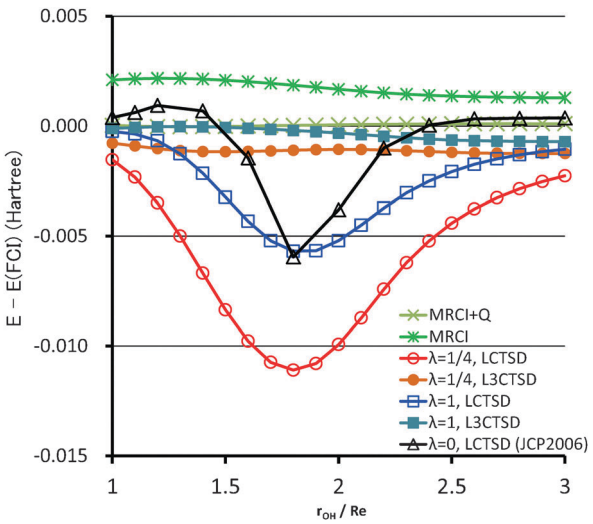

Fig. 1 Potential energy curve errors (from FCI) for the simultaneous bond breaking of H_2O with CASSCF(6e,5o) reference and 6-31G basis sets. L3CTSD is an extension of the LCTSD method using three-body RDMs. Level shift parameters $\lambda = 1/4$ and 1 were tested, while no level shift (*i.e.* $\lambda = 0$) was used in the previous calculations (ref. 109).

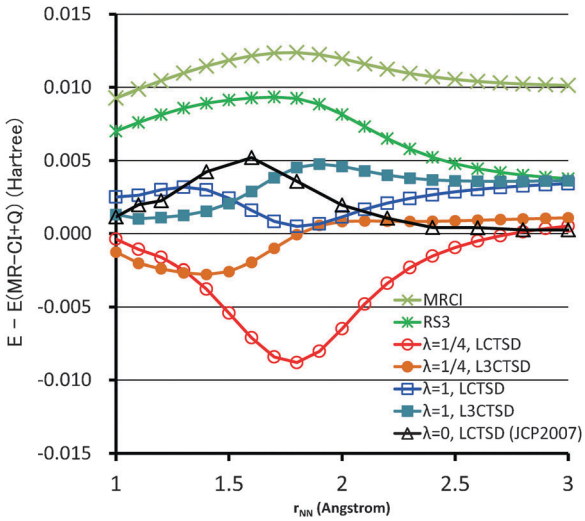

Fig. 2 Potential energy curve errors (from MRIC+Q) for the bond breaking of N_2 with CASSCF(6e,6o) reference and cc-pVDZ basis sets. Level shift parameters $\lambda = 1/4$ and 1 were used. No level shift (*i.e.* $\lambda = 0$) was adopted in the previous calculations (ref. 110), which used the linear dependency threshold $\tau_s = 10^{-1}$ and $\tau_d = 10^{-2}$.

calculations at the several points across the dissociation curves. The plots include the MRCI+Q and CASPT3 results obtained using the MOLPRO program package.

In the CT calculations, we tested two level shift values (1/4 and 1) for λ. With the level-shifted CT conditions, the excessive elimination of amplitude basis was avoided in the overlap orthogonalization [eqn (12)], in which the truncation threshold was then set to 10^{-8} for the removal of linear dependencies. As can be seen in Fig. 1 and 2, the linearized CT singles and doubles (LCTSD) calculations with the level shifts yielded the smooth potential curves. The previous CT calculations without the level shift (*i.e.* with $\lambda = 0$) were performed by setting the truncation thresholds to $t_s = 10^{-1}$ and $t_d = 10^{-2}$ for the semi-internal and doubly-external excitations, respectively.[110] Such an aggressive truncation was required to stabilize the iterative solution of the CT equation, whereas it caused stepwise fluctuations in the dimension of the untruncated basis along the coordinates, resulting in the discontinuity in the curve. Despite the improvements in the continuity of the potentials, the level-shifted LCTSD calculations produced the nonparallelity errors (NPEs) and mean absolute errors (MAEs) more or less similar to those of the previous calculations. The NPEs of the LCTSD curves for H_2O were 9.6, 5.4, and 6.9 mE_h for $\lambda = 1/4$, 1 (with all linearly independent basis), and 0 (with large basis truncation), respectively. With these level shifts, the L3CTSD approach,[112,113] a variant of LCTSD using the connected three-body RDMs instead of neglecting three-body cumulants, provided a significant reduction of the NPEs and MAEs in the potential profiles. This suggests the importance of the connected three-body description in the intermediate of dissociation where the balanced treatment of dynamic and static correlations is critical.

The algorithm described in the previous sections has been implemented into a parallel program, which adopted the Message Passing Interface (MPI), Open Multi-Processing (OpenMP) multi-threading, and the GA toolkit[125] as the parallelization libraries.

We examined the parallel efficiency of our implementation with a DMRG-CT calculation on the perylene molecule, using a large CAS reference correlating 20 π electrons in 20 out-of-plane 2p orbitals of the C atoms, namely CAS(20e,20o), treated by the active-space DMRG method. The wall times were measured with the 6-31G basis set (totally 204 basis functions) and the molecular point group symmetry set to C_1. The first iteration of the CT calculation was performed on a cluster of Linux personal computers (PCs) each with 3.16 GHz Intel Core Duo CPU E8500 (dual cores), 16 GB RAM, and Gigabit Ethernet interface connected through a network switch. Table 4 shows the timings and speedup ratios observed with 4, 10, 20, 30, and 40 CPU cores for the first iteration and its part for evaluating the BCH expansion [eqn (16)]. The parallel scaling was observed to be satisfactory, but the scalability was degraded by the network communication across more processors to access the distributed data of amplitudes through the GA library. The low-latency, high-speed network environments, *e.g.* InfiniBand interconnection, which are nowadays widely available in massively parallel computer systems, are necessary for making an effective use of the GA functions. The evaluation of the BCH expansion to the tenth order occupied 17–18% of the total computation time of the first iteration. The rest of the time was spent mostly on the iterative solution of the amplitude CT equation [eqn (25)].

We have evaluated the vertical excited energies of perylene of 1^1B_{3u} and 2^1A_g electronic states using the cc-pVDZ basis set (340 atomic orbital functions) with the CAS(20e,20o) reference and the geometry optimized with CAM-B3LYP/6-31G* in C_{2v} symmetry. The state-specific DMRG-CASSCF calculations employing 1024 DMRG states were performed to obtain each starting reference of 1^1A_g (ground state), 1^1B_{3u}, and 2^1A_g states. The evaluation of the density matrices of these references was followed by the DMRG-L3CTSD calculations with the level shift $\lambda = 0.1$ and the C 1s orbitals uncorrelated. The excited energies are shown in Table 5 including those of the EOM-CCSD calculations as well as the literature value of the 1^1B_{3u} excited energy for the BLYP/6-31+G* level of theory and experiment.[126] The EOM-CCSD calculations were performed using the tce module[127–129] in the NWCHEM program package.[130] For the 1^1B_{3u} state, the CT method greatly improved the DMRG-CASSCF energy with the inclusion of dynamic correlation, reproducing the EOM-CCSD result based on the HF reference. This indicates that the accuracy of the description for this excited state is dominantly characterized by the dynamic correlation. DMRG-L3CTSD and EOM-CCSD both overestimated the 1^1B_{3u} energy by approximately 0.7 eV, while BLYP/6-31+G* underestimated it by approximately 0.5 eV. The 2^1A_g state was in contrast found to be a multireference state in the sense that the DMRG-L3CTSD and EOM-CCSD predictions dissociated. Because the 2^1A_g energies obtained with DMRG-CASSCF and DMRG-L3CTSD were similar, the dynamic correlation did not seem to be so important in this state. The 2^1A_g state is a symmetrically forbidden state, so that it has not been measured experimentally. Our calculations predicted that there is the dark 2^1A_g state higher lying at 0.7 eV above the visible 1^1B_{3u} state.

Table 4 Wall times (in seconds) and parallel scaling ratios (in parentheses) for the first iteration and BCH expansion in the L3CTSD calculation on perylene/6-31G in C1 symmetry. The total time includes the time for the BCH expansion. Two threads were assigned to every two CPU cores, which correspond to a single PC node. Parallel efficiencies are shown in percent

		The first iteration			
		Total		BCH expansion	
Number of CPU cores	Ratio	Time/sec	Ratio	Time/sec	Ratio
4	(1.0)	218 640	(1.0) 100%	39 398	(1.0) 100%
10	(2.5)	97 819	(2.2) 89%	17 948	(2.2) 88%
20	(5.0)	54 303	(4.0) 81%	9 682	(4.1) 81%
30	(7.5)	40 202	(5.4) 73%	7 009	(5.6) 75%
40	(10.0)	33 347	(6.5) 66%	5 328	(7.4) 74%

Table 5 Vertical excitation energies (in eV) of the 1^1B_{3u} and 2^1A_g states for perylene. The DMRG-CASSCF, DMRG-L3CTSD, and EOM-CCSD calculations were performed using the cc-pVDZ basis set

Method	1^1B_{3u}	2^1A_g
DMRG-CASSCF(20e,20o)	4.89	4.38
DMRG-L3CTSD	3.65	4.35
EOM-CCSD	3.64	5.16
BLYP/6-31+G* [a]	2.49	
exptl.[a]	2.96	

[a] Ref. 126. The 2^1A_g state is a forbidden excitation.

5 Summary

In the previous work we proposed the combination of DMRG and CT methods to compute a good description of the multi-reference problems which require a high-order treatment of the dynamic correlation as well as the ability to treat large active space. The key computational advance that has overcome the high expense of CT calculations using large-size basis sets for recovering a large amount of dynamic correlation is the parallelized algorithm and its computer implementation of the CT theory. This paper has provided a detailed description of our parallelization as well as the tensor contraction expressions. A highlight of the parallelization was the way of distributing the storage of four-index arrays for two-body Hamiltonian elements and amplitudes across processors. The second focus of this study was on stabilizing the iterative CT solution by a modification to the amplitude equation with the inclusion of the level shift parameter. The level-shifted condition numerically removes the intruder states easily so that all linearly independent orthogonal basis states are incorporated with the small truncation into the amplitude and thus the resultant potential energy curves maintained continuity. This way of modifying the stationary equation can be applied to the other MRCC-type methods, which also often have convergence difficulty in the solution. Although the parallelization speeds up CT calculations, further investigations to reduce the total costs of LCTSD calculations by approximation are necessary for challenging application to much larger multireference systems.

Acknowledgements

This research was supported in part by Grant-in-Aid for Scientific Research (C) (Grant No. 21550027) and Young Scientists (B)

(Grant No. 21750028) from Ministry of Education, Culture, Sports, Science and Technology-Japan (MEXT).

References

1 B. O. Roos, P. Linse, P. E. M. Siegbahn and M. R. A. Blomberg, *Chem. Phys.*, 1982, **66**, 197.
2 K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej and K. Wolinski, *J. Phys. Chem.*, 1990, **94**, 5483.
3 K. Andersson, P.-Å. Malmqvist and B. O. Roos, *J. Chem. Phys.*, 1992, **96**, 1218.
4 J. Finley, P.-Å. Malmqvist, B. O. Roos and L. Serrano-Andrés, *Chem. Phys. Lett.*, 1998, **288**, 299.
5 F. Aquilante, P.-Å. Malmqvist, T. B. Pedersen, A. Ghosh and B. O. Roos, *J. Chem. Theory Comput.*, 2008, **4**, 694.
6 P.-Å. Malmqvist, K. Pierloot, A. R. M. Shahi, C. J. Cramer and L. Gagliardi, *J. Chem. Phys.*, 2008, **128**, 204109.
7 K. Wolinski, H. L. Sellers and P. Pulay, *Chem. Phys. Lett.*, 1987, **140**, 225.
8 K. Wolinski and P. Pulay, *J. Chem. Phys.*, 1989, **90**, 3647.
9 P. Pulay, *Int. J. Quantum Chem.*, 2011, **111**, 3273.
10 J. J. W. McDouall, K. Peasley and M. A. Robb, *Chem. Phys. Lett.*, 1988, **148**, 183.
11 R. B. Murphy and R. P. Messmer, *Chem. Phys. Lett.*, 1991, **183**, 443.
12 R. B. Murphy and R. P. Messmer, *J. Chem. Phys.*, 1992, **97**, 4170.
13 K. Hirao, *Chem. Phys. Lett.*, 1992, **190**, 374.
14 K. Hirao, *Chem. Phys. Lett.*, 1992, **196**, 397.
15 H. Nakano, *J. Chem. Phys.*, 1993, **99**, 7983.
16 H. Nakano, *Chem. Phys. Lett.*, 1993, **207**, 372.
17 H. Nakano, J. Nakatani and K. Hirao, *J. Chem. Phys.*, 2001, **114**, 1133.
18 H. Nakano, R. Uchiyama and K. Hirao, *J. Comput. Chem.*, 2002, **23**, 1166.
19 M. Miyajima, Y. Watanabe and H. Nakano, *J. Chem. Phys.*, 2006, **124**, 044101.
20 S. Ten-no, *Chem. Phys. Lett.*, 2007, **447**, 175.
21 W. Mizukami, Y. Kurashige, M. Ehara, T. Yanai and T. Itoh, *J. Chem. Phys.*, 2009, **131**, 174313.
22 H.-J. Werner, *Mol. Phys.*, 1996, **89**, 645.
23 P. Celani and H.-J. Werner, *J. Chem. Phys.*, 2000, **112**, 5546.
24 P. Celani and H.-J. Werner, *J. Chem. Phys.*, 2003, **119**, 5044.
25 T. Shiozaki and H.-J. Werner, *J. Chem. Phys.*, 2010, **133**, 141103.
26 T. Shiozaki, W. Győrffy, P. Celani and H.-J. Werner, *J. Chem. Phys.*, 2011, **135**, 081106.
27 C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger and J. P. Malrieu, *J. Chem. Phys.*, 2001, **114**, 10252.
28 C. Angeli, R. Cimiraglia and J.-P. Malrieu, *J. Chem. Phys.*, 2002, **117**, 9138.
29 C. Angeli, B. Bories, A. Cavallini and R. Cimiraglia, *J. Chem. Phys.*, 2006, **124**, 054108.
30 C. Angeli, M. Pastore and R. Cimiraglia, *Theor. Chem. Acc.*, 2007, **117**, 743.
31 Y. Kurashige and T. Yanai, *J. Chem. Phys.*, 2011, **135**, 094104.
32 R. J. Buenker and S. D. Peyerimhoff, *Theor. Chim. Acta*, 1968, **12**, 183.
33 R. J. Buenker, S. D. Peyerimhoff and W. Butscher, *Mol. Phys.*, 1978, **35**, 771.
34 B. Roos, *Chem. Phys. Lett.*, 1972, **15**, 153.
35 R. J. Buenker and S. D. Peyerimhoff, *Theor. Chim. Acta*, 1974, **35**, 33.
36 W. Meyer, in *Methods of electronic structure theory*, ed. H. F. Schaefer III, Plenum, New York, 1977, pp. 413–446.
37 E. R. Davidson and D. W. Silver, *Chem. Phys. Lett.*, 1977, **52**, 403.
38 R. Gdanitz and R. Ahlrichs, *Chem. Phys. Lett.*, 1988, **143**, 413.
39 P. G. Szalay and R. J. Bartlett, *Chem. Phys. Lett.*, 1993, **214**, 481.
40 M. Torheyden and E. Valeev, *J. Chem. Phys.*, 2009, **131**, 171103.
41 P. E. M. Siegbahn, *J. Chem. Phys.*, 1980, **72**, 1647.
42 M. Kleinschmidt, J. r. Tatchen and C. M. Marian, *J. Comput. Chem.*, 2002, **23**, 824.
43 T. Müller, *J. Phys. Chem. A*, 2009, **113**, 12729.
44 H.-J. Werner and E. Reinsch, *J. Chem. Phys.*, 1982, **76**, 3144.
45 H.-J. Werner and P. J. Knowles, *J. Chem. Phys.*, 1988, **89**, 5803.
46 P. J. Knowles and H.-J. Werner, *Chem. Phys. Lett.*, 1988, **145**, 514.
47 T. Shiozaki, G. Knizia and H.-J. Werner, *J. Chem. Phys.*, 2011, **134**, 034113.
48 K. R. Shamasundar, G. Knizia and H.-J. Werner, *J. Chem. Phys.*, 2011, **135**, 054101.
49 T. Shiozaki and H.-J. Werner, *J. Chem. Phys.*, 2011, **134**, 184104.
50 A. Banerjee and J. Simons, *Int. J. Quantum Chem.*, 1981, **19**, 207.
51 A. Banerjee, *J. Chem. Phys.*, 1982, **76**, 4548.
52 S. Pal, M. Durga Prasad and D. Mukherjee, *Theor. Chim. Acta*, 1983, **62**, 523.
53 A. Banerjee and J. Simons, *Chem. Phys.*, 1983, **81**, 297.
54 S. Pal, *Theor. Chim. Acta*, 1984, **66**, 207.
55 A. Banerjee and J. Simons, *Chem. Phys.*, 1984, **87**, 215.
56 N. Oliphant and L. Adamowicz, *J. Chem. Phys.*, 1991, **94**, 1229.
57 M. R. Hoffmann and J. Simons, *J. Chem. Phys.*, 1988, **88**, 993.
58 W. D. Laidig, P. Saxe and R. J. Bartlett, *J. Chem. Phys.*, 1987, **86**, 887.
59 R. J. Bartlett and J. Noga, *Chem. Phys. Lett.*, 1988, **150**, 29.
60 R. J. Bartlett, S. A. Kucharski and J. Noga, *Chem. Phys. Lett.*, 1989, **155**, 133.
61 J. D. Watts, G. W. Trucks and R. J. Bartlett, *Chem. Phys. Lett.*, 1989, **157**, 359.
62 S. R. White, *J. Chem. Phys.*, 2002, **117**, 7472.
63 T. Kinoshita, O. Hino and R. J. Bartlett, *J. Chem. Phys.*, 2005, **123**, 074106.
64 O. Hino, T. Kinoshita, G. K.-L. Chan and R. J. Bartlett, *J. Chem. Phys.*, 2006, **124**, 114311.
65 M. Kállay, P. G. Szalay and P. R. Surján, *J. Chem. Phys.*, 2002, **117**, 980.
66 J. A. Parkhill and M. Head-Gordon, *J. Chem. Phys.*, 2010, **133**, 124102.
67 S. Li, *J. Chem. Phys.*, 2004, **120**, 5017.
68 F. A. Evangelista and J. Gauss, *J. Chem. Phys.*, 2011, **134**, 114102.
69 F. A. Evangelista and J. Gauss, *Chem. Phys.*, 2011, published online.
70 M. Hanauer and A. Köhn, *J. Chem. Phys.*, 2011, **134**, 204111.
71 D. Datta, L. Kong and M. Nooijen, *J. Chem. Phys.*, 2011, **134**, 214116.
72 B. O. Roos, P. R. Taylor and P. E. M. Si gbahn, *Chem. Phys.*, 1980, **48**, 157.
73 B. O. Roos, *Adv. Chem. Phys.*, 1987, **69**, 399.
74 K. Ruedenberg, M. W. Schmidt, M. M. Gilbert and S. T. Elbert, *Chem. Phys.*, 1982, **71**, 41.
75 S. R. White, *Phys. Rev. Lett.*, 1992, **69**, 2863.
76 S. R. White, *Phys. Rev. B: Condens. Matter*, 1993, **48**, 10345.
77 S. R. White and R. L. Martin, *J. Chem. Phys.*, 1999, **110**, 4127.
78 S. Daul, I. Ciofini, C. Daul and S. R. White, *Int. J. Quantum Chem.*, 2000, **79**, 331.
79 A. O. Mitrushenkov, G. Fano, F. Ortolani, R. Linguerri and P. Palmieri, *J. Chem. Phys.*, 2001, **115**, 6815.
80 G. K.-L. Chan and M. Head-Gordon, *J. Chem. Phys.*, 2002, **116**, 4462.
81 G. K.-L. Chan and M. Head-Gordon, *J. Chem. Phys.*, 2003, **118**, 8551.
82 O. Legeza, J. Röder and B. Hess, *Phys. Rev. B: Condens. Matter*, 2003, **67**, 125114.
83 A. O. Mitrushenkov, R. Linguerri, P. Palmieri and G. Fano, *J. Chem. Phys.*, 2003, **119**, 4148.
84 G. K.-L. Chan, *J. Chem. Phys.*, 2004, **120**, 3172.
85 G. K.-L. Chan, M. Kállay and J. Gauss, *J. Chem. Phys.*, 2004, **121**, 6110.
86 G. K.-L. Chan and T. Van Voorhis, *J. Chem. Phys.*, 2005, **122**, 204101.
87 G. Moritz, B. A. Hess and M. Reiher, *J. Chem. Phys.*, 2005, **122**, 024107.
88 G. Moritz, A. Wolf and M. Reiher, *J. Chem. Phys.*, 2005, **123**, 184105.
89 J. Hachmann, W. Cardoen and G. K.-L. Chan, *J. Chem. Phys.*, 2006, **125**, 144101.
90 G. Moritz and M. Reiher, *J. Chem. Phys.*, 2006, **124**, 034103.
91 J. J. Dorando, J. Hachmann and G. K.-L. Chan, *J. Chem. Phys.*, 2007, **127**, 084109.
92 G. Moritz and M. Reiher, *J. Chem. Phys.*, 2007, **126**, 244109.

93 G. K.-L. Chan, *Phys. Chem. Chem. Phys.*, 2008, **10**, 3454.
94 G. K.-L. Chan, J. J. Dorando, D. Ghosh, J. Hachmann, E. Neuscamman, H. Wang and T. Yanai, *Frontiers in Quantum Systems in Chemistry and Physics*, Springer Netherlands, Dordrecht, 2008, vol. 18, p. 49.
95 K. H. Marti, I. M. Ondk, G. Moritz and M. Reiher, *J. Chem. Phys.*, 2008, **128**, 014104.
96 D. Zgid and M. Nooijen, *J. Chem. Phys.*, 2008, **128**, 014107.
97 D. Zgid and M. Nooijen, *J. Chem. Phys.*, 2008, **128**, 144115.
98 D. Zgid and M. Nooijen, *J. Chem. Phys.*, 2008, **128**, 144116.
99 D. Ghosh, J. Hachmann, T. Yanai and G. K.-L. Chan, *J. Chem. Phys.*, 2008, **128**, 144117.
100 T. Yanai, Y. Kurashige, D. Ghosh and G. K.-L. Chan, *Int. J. Quantum Chem.*, 2009, **109**, 2178–2190.
101 J. J. Dorando, J. Hachmann and G. K.-L. Chan, *J. Chem. Phys.*, 2009, **130**, 184111.
102 G. K.-L. Chan and D. Zgid, *Annu. Rep. Comput. Chem.*, 2009, **5**, 149.
103 Y. Kurashige and T. Yanai, *J. Chem. Phys.*, 2009, **130**, 234114.
104 K. Marti and M. Reiher, *Mol. Phys.*, 2010, **108**, 501.
105 W. Mizukami, Y. Kurashige and T. Yanai, *J. Chem. Phys.*, 2010, **133**, 091101.
106 G. K.-L. Chan and S. Sharma, *Annu. Rev. Phys. Chem.*, 2011, **62**, 465.
107 T. Yanai, Y. Kurashige, E. Neuscamman and G. K.-L. Chan, *J. Chem. Phys.*, 2010, **132**, 024105.
108 E. Neuscamman, T. Yanai and G. K.-L. Chan, *J. Chem. Phys.*, 2010, **132**, 024106.
109 T. Yanai and G. K.-L. Chan, *J. Chem. Phys.*, 2006, **124**, 194106.
110 T. Yanai and G. K.-L. Chan, *J. Chem. Phys.*, 2007, **127**, 104107.
111 G. K.-L. Chan and T. Yanai, *Adv. Chem. Phys.*, 2007, **134**, 343.
112 E. Neuscamman, T. Yanai and G. K.-L. Chan, *J. Chem. Phys.*, 2009, **130**, 124102.
113 E. Neuscamman, T. Yanai and G. K.-L. Chan, *J. Chem. Phys.*, 2009, **130**, 169901.
114 E. Neuscamman, T. Yanai and G. K.-L. Chan, *Int. Rev. Phys. Chem.*, 2010, **29**, 231.
115 W. Kutzelnigg and D. Mukherjee, *J. Chem. Phys.*, 1997, **107**, 432.
116 D. Mukherjee, *Chem. Phys. Lett.*, 1997, **274**, 561.
117 D. Mukherjee, *Recent Progress in Many-Body Theories*, Plenum, New York, 1995, vol. 4, p. 127.
118 W. Kutzelnigg, *Chem. Phys. Lett.*, 1979, **64**, 383.
119 B. O. Roos and K. Andersson, *Chem. Phys. Lett.*, 1995, **245**, 215.
120 K. Andersson, *Theor. Chim. Acta*, 1995, **91**, 31.
121 G. Ghigo, B. O. Roos and P.-Å. Malmqvist, *Chem. Phys. Lett.*, 2004, **396**, 142.
122 N. Forsberg and P.-Å. Malmqvist, *Chem. Phys. Lett.*, 1997, **274**, 196.
123 Y. Choe, H. Witek, J. Finley and K. Hirao, *J. Chem. Phys.*, 2001, **114**, 3913.
124 H. Witek, Y. Choe, J. Finley and K. Hirao, *J. Comput. Chem.*, 2002, **23**, 957–965.
125 J. Nieplocha, R. Harrison and R. Littlefield, *J. Supercomput.*, 1996, **10**, 169.
126 T. Halasinski, J. Weisman, R. Ruiterkamp, T. Lee, F. Salama and M. Head-Gordon, *J. Phys. Chem. A*, 2003, **107**, 3660.
127 S. Hirata, *J. Phys. Chem. A*, 2003, **107**, 9887.
128 S. Hirata, *J. Chem. Phys.*, 2004, **121**, 51.
129 S. Hirata, *Theor. Chem. Acc.*, 2006, **116**, 2.
130 M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus and W. A. de Jong, *Comput. Phys. Commun.*, 2010, **181**, 1477.

